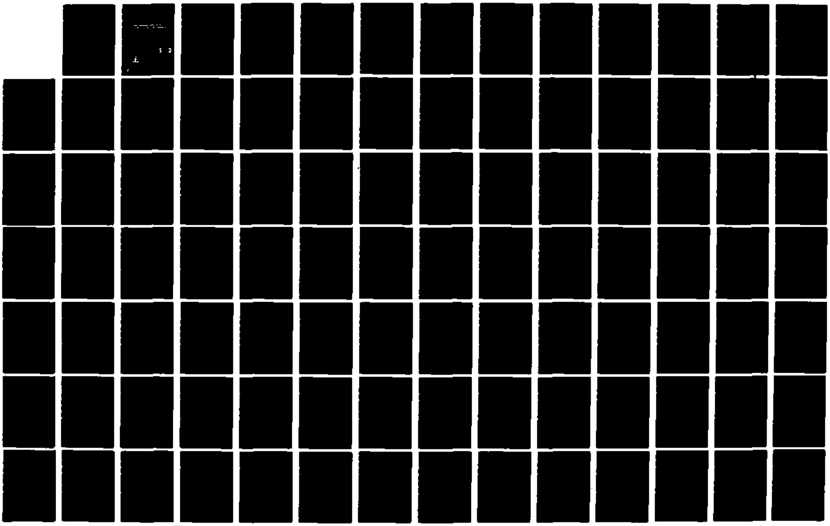
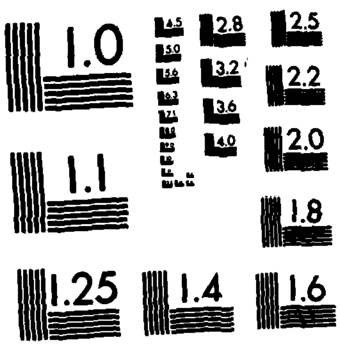
AD-A125 337 A MIXED REGRESSION MODEL FOR THE ANALYSIS OF EXPLOSIVES PERFORMANCE DATA(U) NAVAL SURFACE WEAPONS CENTER SILVER SPRING MD W W MCDONALD 30 JUL 82 NSWC/TR-82-74 1/2

UNCLASSIFIED SBI-AD-F500 124 F/G 19/4 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

AD A125337

NSWC TR 82-74

(12)

AD-F 600 124

A MIXED REGRESSION MODEL FOR THE ANALYSIS OF EXPLOSIVES PERFORMANCE DATA

BY W. MCDONALD

RESEARCH AND TECHNOLOGY DEPARTMENT

30 JULY 1982

NAVAL SURFACE WEAPONS CENTER

Dahlgren, Virginia 22448 • Silver Spring, Maryland 20910

83 02 017 026

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NSWC TR 82-74	2. GOVT ACCESSION NO.: AD-A125 337
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): A MIXED REGRESSION MODEL FOR THE ANALYSIS OF EXPLOSIVES PERFORMANCE DATA	5. TYPE OF REPORT & PERIOD COVERED: Interim
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): William W. McDonald	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Surface Weapons Center (Code R14) White Oak Silver Spring, MD 20910	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62633N; F33354 SF33354391; R10BA
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE: 30 July 1982
	13. NUMBER OF PAGES: 98
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Explosives performance
Shock wave measurements
Statistics
Maximum likelihood
Mixed model
Regression
Variance components
Restricted maximum likelihood

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An intra-class regression model with additive error terms consisting of a gage calibration error, a within-class performance variation, and a general experimental error is developed for the analysis of explosives performance data associated with the shock waves of underwater explosions. Parameter estimates for this mixed model are sought by means of the maximum likelihood and restricted maximum likelihood techniques. Derivatives of the likelihood function suitable for constrained Newton-Raphson optimization are derived. Useful applications are discussed.

FOREWORD

A statistical methodology is presented for the analysis of explosives performance data associated with the shock waves of underwater explosions. The model should increase both the accuracy and amount of useful information now being extracted from the data. In addition, it should serve as a basic tool for the development of statistical tests for the comparisons of explosives and for the study of more efficient experimental designs. The work was funded through the Explosives Development, Effects and Safety Block of the Naval Sea Systems Command (Task Area SF-33-354-391) as a part of the MADAM program.

Approved by:

J. F. PROCTOR, Head

Energetic Materials Division

CONTENTS

CONTENTS (Cont.)

ILLUSTRATIONS

TABLES

CHAPTER 1

INTRODUCTION

In this report we develop a general statistical model for the analysis of explosives performance data. The data of interest consist of observations of an arbitrary measure of performance that can be derived from the transient pressure-induced responses of piezo-electric gages located in the vicinities of underwater explosions. Applications of the model concern the variations of performance within broadly defined classes of explosive charges and the predictions and comparisons, at specified ranges, of the performances of charges belonging to these classes. The observations are complicated by the presence of gage calibration errors and various errors that arise from the measurement and data processing techniques employed.

The traditional and current method for analyzing data of this kind (see Cole, Reference 1, p. 240) is to express the relationship between the scaled performance variable and the scaled distance from the charge as a power law whose parameters are determined from an ordinary least squares fit of a straight line to the logarithmically transformed data. Relationships obtained in this manner are popularly called "similitude equations" from the principle of similarity that underlies the scaling of explosion shock wave phenomena. The

[1]Cole, R. H., Underwater Explosions (New Jersey: Princeton Univ. Press, 1948).

usual method of scaling is called Hopkinson or cube-root scaling (see e.g. Snay, Reference 2), which refers to the fact that the time and length scale factors are proportional to the cube root of the charge weight (or any other proportional measure of the explosion energy). Recently Goertner (Reference 3) has extended this scaling method to include variations with ambient water sound speeds and densities. It is now customary to compile similitude equations for a large variety of explosive classes for the following measures of performance: the peak pressure, a characteristic time constant, and the impulses and energies per unit area delivered to a given location by the shock wave within various multiples of the time constant.

The application of more sophisticated statistical techniques to this area has been blocked, perhaps primarily, by complicated dependencies within the data. Correlations exist among the observations obtained from a single test, as pointed out by Brown (Reference 4), and among observations obtained with the same gages and gage calibration constants. The mixed linear models necessary for an adequate statistical treatment of such data constitute an active area of current research.

The model that we develop below generalizes the current approach by accepting arbitrary measures of performance and allowing regression functions of arbitrary form that are linear in their coefficients (such as higher degree

[2]Snay, H. G., "Model Tests and Scaling," NOLTR 63-257, 1 Dec 1964.

[3]Goertner, J. F., "Scaling Underwater Explosion Shock Waves for Differences in Ambient Sound Speed and Density," NSWC TR 80-491, 18 Dec 1980.

[4]Brown, R. H., "Analysis of Data When Several Sources of Variation are Present," Explosives Research Memorandum 22, Navy Dept. Bureau of Ordnance, 1 Dec 1944.

polynomials). As the principle of shock wave similarity is still adhered to we continue to refer to the equation of the regression mean as the similitude equation for the particular explosive class--a term that we will use to emphasize the fact that explosive charges are more properly viewed as members of a class of objects that in many respects are alike but which differ in ways that affect the observed measures of performance.

The model extends the presently used approach by explicitly including sources of random variation in its formulation. In the interest of model simplicity our philosophy has been to include only those sources that are thought to produce significantly large effects and are unavoidable. Complicating effects that are avoidable or correctible will be presumed to have been eliminated either by an appropriate reprocessing of the data or by modifications of the experimental techniques.

A possible model deficiency is that no explicit treatment of so called batch effects is included. This refers to well recognized performance variations among charges taken from different batches or preparations of the same explosive material. It was felt that batch effects did not justify the further complication of an already complex model and that they could be handled in another manner such as by treating different batches as different explosive classes, by increasing the number of batches and randomizing the charge selection, and by improving the explosive preparation quality control. In any case the use of the model should be made with the possibility of batch effects borne in mind, and a thorough examination of the model residuals for the presence of these and any other systematic effects is recommended (see Section 5-1).

The report is divided into five chapters. Chapter 2 discusses the development of the model in detail. The model may be described as an intra-class regression model with an additive error term consisting of a gage class calibration error, a within-class performance variation, and a general experimental error. Under a suitable transformation of the performance variable, multivariate normality of the errors is assumed. Chapter 3 discusses the estimation of the model parameters by the methods of maximum likelihood and restricted maximum likelihood. The derivation of derivatives needed for an iterative solution of the likelihood equations follows the approach of Harville (Reference 5). In Chapter 4 we find a brief description of both unconstrained and constrained Newton-Raphson and method of scoring optimization techniques and related topics. And finally in Chapter 5 we give some practical applications of the model.

---

[5] Harville, D. A., "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," Journal of the American Statistical Assoc., Vol. 72, No 358, 1977, p. 320.

CHAPTER 2

MODEL DEVELOPMENT

We will denote the response and regressor variables of the model as y and x respectively, often with subscripts to specify a particular observation. Vectors in the model will be indicated by use of the underbar notation; hence, a sample of response variables will appear as $\underline{y}$. No notational distinction will be made between realized and random samples, but this difference should be apparent from the context.

In the theory of linear models it is common to deal with transformed response and regressor variables to promote variance homogeneity, model simplicity, and other desirable model properties. Thus, we define y as a possibly transformed value of a scaled measure of performance, as discussed earlier, and x as a possibly transformed scaled distance taken to be zero at the charge center. This is in accord with past derivations of shock wave similitude equations in which straight lines are fitted to the logarithms of the scaled data. Hence, the present model will be compatible with these forms.

Development of the model will be based upon a number of reasonable assumptions which will appear throughout this section. To these and the assumption of shock wave similarity already made we add that the water between the charge and gages is assumed t' be homogeneous so that disturbances are propaged through the water in a regular manner, and we assume the values of x

to be accurately known. Recent tests of the last assumption have verified its accuracy when distances are determined from the measured times of arrival (Reference 6).

2-1 BASIC MODEL FOR A SINGLE OBSERVATION

In the sample of y values let y_{ijkm} be the observation made in the jth shot of the ith explosive class with a gage of the kth gage class having calibration index m. The gage is located at the (transformed, scaled) distance x_{ijkm}. The gage class index refers to one of several broad classes of gages such as 1/4 inch tourmaline, 3/8 inch tourmaline, 1/2 inch tourmaline etc. In a typical test it is common to employ a string of perhaps 10 to 12 gages placed so that smaller diameter gages are grouped closer to the charge and larger diameter gages grouped farther from the charge. Typically gages from 3 to 4 gage classes are used. A gage is usually recalibrated prior to each test program and assigned a gage constant (units of picocoulombs/psi) which is used to calculate all pressures measured by the gage until it is recalibrated. From this it is clear that measures of performance derived from the various pressure-time records of a particular gage could be correlated, i.e., all affected in the same manner. During the lifetime of a gage, lasting only a single shot to perhaps several years, it might be recalibrated as many as 5 to 10 times.

The system by which observations will be indexed is described as follows. Let C and K be the total numbers of explosive classes and gage classes of

[6]Gaspin, J. B., "Validation of a Gage Location Method for Underwater Explosion Tests," NSWC TR in preparation.

interest, respectively, and let J_i be the total number of shots of the ith explosive class. Then index i will simply run from 1 to C, index j will run from 1 to J_i, and index k will run from 1 to K. The numbering assigned to particular classes or shots is arbitrary. In the same sense the gage calibration index m is also arbitrarily assigned, but it will require a somewhat lengthier explanation. Basically, we wish to identify those observations obtained from the same gage and one of its particular gage calibration constants and assign to these observations the same value of m. We will establish separate sets of m values for the observations associated with each gage class of interest. For the kth gage class, the total number M_k of such m values is equal to the total number of unique ordered pairs (n,d) among the observations associated with the kth gage class, where n is a unique identifying number of the gage and d is the date on which the calibration session was conducted. hence, for the kth gage class the calibration index m will run from 1 to M_k. Note that, although several measurements can have the same indices k and m, an observation is uniquely labeled by the set of indices (i,j,k,m), since only a single observation can be obtained from a particular gage on a particular shot.

Consider, now, the jth test of the ith explosive class. We assume that the passage of the shock wave through the water induces a particular unknown functional relationship between the quantity we are seeking to measure and the travel distance x. We will denote a representation of this function as $f_{ij}(x)$ for the jth test of the ith explosive class. We can write the observation obtained with a gage having indices k and m as

$$y_{ijkm} = f_{ij}(x_{ijkm}) + e_{ijkm}, \tag{2-1}$$

where e_{ijkm} is the deviation of the measured value from the true value as indicated in Figure 1. Values of y measured by other gages and at other values of x will be scattered about $f_{ij}(x)$ in some manner that we will now consider.

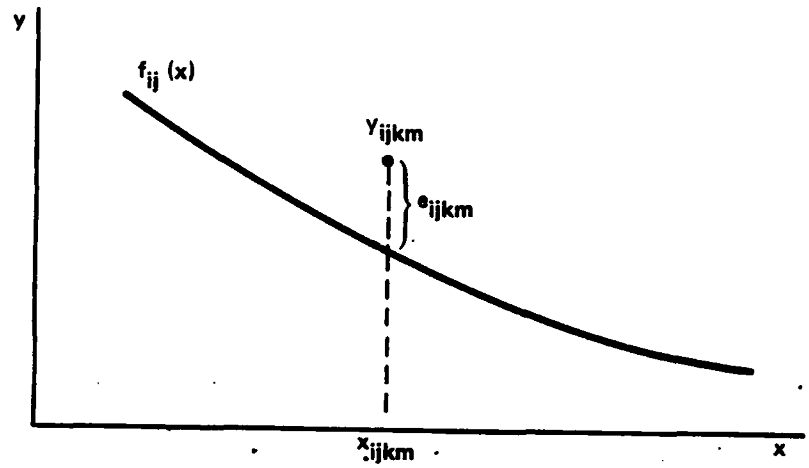


Figure 1. Response y versus distance x for jth shot

We will assume the error e_{ijkm} to be a random variable with a mean of zero. This requires that any bias in e_{ijkm} introduced, for example, by an incorrect gage size correction or other systematic measurement effect, be removed prior to the statistical analysis. Current thinking by experimentalists is that this is not an unreasonable assumption. Furthermore, there are two compelling reasons for it from theoretical grounds. First, it keeps the model from becoming unmanageably complex. Second, and most importantly, the presence of bias terms would make the usual similitude equations unestimable. That is, they would have no unique solutions. An explanation for this is given below in Section 3-1. Thus, in the model we require

$$E(e_{ijkm}) = 0. \tag{2-2}$$

We will assume that e_{ijkm} can, however, be separated into two additive random parts as

$$e_{ijkm} = \alpha^*_{km} + \epsilon^*_{ijkm} \; , \tag{2-3}$$

each with zero mean. The first of these, α^*_{km}, will denote that part that is due to gage calibration error. We will assume that all gages of a particular gage class have the same distribution of calibration errors. For this reason a previously used but recalibrated gage will be treated in the same manner as a new gage--the calibration errors of both are drawn from the same distribution. The remaining component in (2-3), ϵ^*_{ijkm}, is regarded as a general error term arising from the measurement and data handling processes but from no source in particular. This decomposition of e_{ijkm} is illustrated in Figure 2. Substitution of (2-3) into (2-1) yields

$$y_{ijkm} = f_{ij}(x_{ijkm}) + \alpha^*_{km} + \epsilon^*_{ijkm}. \tag{2-4}$$

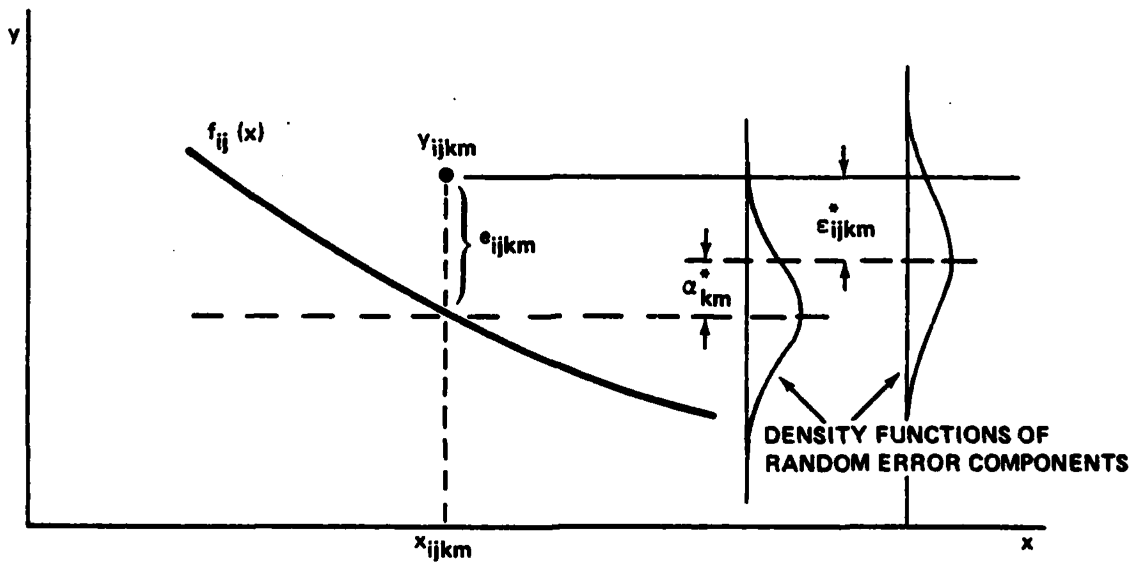


Figure 2. Decomposition of e_{ijkm} for jth shot

It is convenient and consistent with past practices to represent the functional relationship $f_{ij}(x)$ by a polynomial in x. This can be written in a somewhat more general fashion for the jth shot of the ith explosive class as

$$f_{ij}(x_{ijkm}) = \underline{\phi}'_{ijkm}\,\underline{\mu}^+_{ij} \quad , \tag{2-5}$$

where $\underline{\phi}_{ijkm}$ is a vector function of the regressor variable x_{ijkm} (prime denotes the transpose), and $\underline{\mu}^+$ is a vector of unknown parameter values that is conformable with $\underline{\phi}'_{ijkm}$. In the traditional case of similitude equations we have $\underline{\phi}'_{ijkm} = (1, x_{ijkm})$. For the case of a (p-1)th degree polynomial we have, of course, $\underline{\phi}'_{ijkm} = (1, x_{ijkm}, x^2_{ijkm}, \ldots, x^{p-1}_{ijkm})$. Generally, for underwater explosions a low degree polynomial with $p = 2$ or 3 will be adequate. For other applications where the choice of $p > 6$ may seem more appropriate it may be preferable to define $\underline{\phi}_{ijkm}$ in terms of Chebyshev or orthogonal polynomials to avoid problems of ill conditioning (see Seber, Reference 7, p. 214). The generality of expression (2-5) should be fully appreciated. In addition to functions of a single dimension it will admit the use of multidimensional functions (that are linear in their coefficients) as well. In the subsequent development we will assume $\underline{\phi}_{ijkm}$ to consist of p elements which could be any of these types.

The relationship between y and x will vary from one explosive class to another and also between shots within a particular explosive class. Within-class variations are considered to be caused by charge fabrication and preparation methods that are difficult or impossible to control precisely, and

[7]Seber, G. A. F., Linear Regression Analysis (New York: John Wiley & Sons, Inc., 1977).

also by naturally occurring chemical and physical inhomogeneities randomly located within the explosive mixtures. Mathematical models for these relationships must, consequently, be of a random nature. We can view the relationships between y and x generated by the shots of a particular explosive class as a family of curves each of which is similar to the one illustrated in Figure 1. This idea is illustrated in Figure 3.

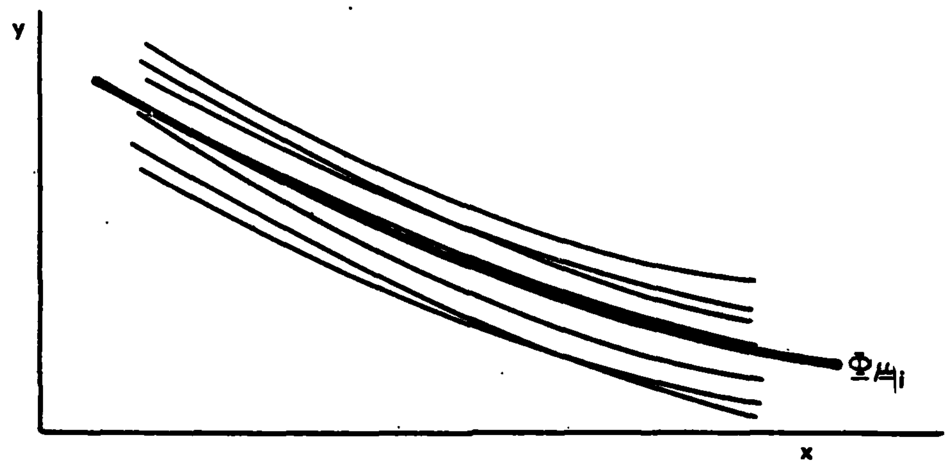


Figure 3. Response versus distance curves for an explosive class

A mathematical model for the behavior of an explosive class can be constructed by allowing the parameters $\underline{\mu}^{+}_{ij}$ of equation (2-5) to be random variables. Mean values, variances, and covariances of the parameters will then be constants of the explosive class. If we denote the mean of $\underline{\mu}^{+}_{ij}$ as $\underline{\mu}_i$, we can then write

$$\underline{\mu}^{+}_{ij} = \underline{\mu}_i + \underline{\beta}^{*}_{ij} , \qquad (2\text{-}6)$$

which decomposes $\underline{\mu}^{+}_{ij}$ into the sum of a vector of fixed effects $\underline{\mu}_i$ and a vector of random effects with zero means $\underline{\beta}^{*}_{ij}$. Inserting equations (2-6) into (2-5) we obtain

$$f_{ij}(x_{ijkm}) = \underline{\phi}'_{ijkm}\,\underline{\mu}_i + \underline{\phi}'_{ijkm}\,\underline{\beta}^*_{ij}. \tag{2-7}$$

Note that $f_{ij}(x)$ is now regarded as a random function consisting of a mean curve and a random variation about that mean.

By inserting (2-7) into equation (2-4) we obtain the basic model for the analysis of explosion test data. It is convenient to write this as

$$y_{ijkm} = \underline{\phi}'_{ijkm}\,\underline{\mu}_i$$

$$+ \alpha^*_{km} + \underline{\phi}'_{ijkm}\,\underline{\beta}^*_{ij} + \epsilon^*_{ijkm}, \tag{2-8}$$

where the second line is the random "error" term with zero mean. The basic model is recognized as a linear regression model with a complicated error structure. It may also be referred to as a mixed model or mixed regression model.

Equation (2-8) expresses the contention that a single observation may be viewed as a realization of a random variation about the mean response of the ith explosive class. Furthermore, it holds that the random part consists of the sum of (1) a random gage calibration term, (2) the random performance variations of ("identical") charges within the explosive class, and (3) a random experimental error term. And finally, it indicates that the mean (or expected value) of the observations of the ith explosive class, that corresponds to what is traditionally called the explosive similitude equation (expressed in terms of the transformed variable y) is given by

$$E(y_{ijkm}) = \underline{\phi}'_{ijkm}\,\underline{\mu}_i \tag{2-9}$$

at distance x_{ijkm}.

2-2 THE MODEL FOR A VECTOR OF OBSERVATIONS

We now express the model in matrix form in three steps by first writing the model for the measurements of a particular shot (i.e., for fixed values of i and j), then combining these to form the model for the ith explosive class, and finally combining several single explosive class models to obtain a model for multiple explosive classes.

To reiterate our previous discussion concerning indices we are interested in a total of C explosive classes and K gage classes. Also we let the number of shots in the ith explosive class be J_i and the number of gage calibrations in the kth gage class be M_k. Hence, the indices take on the values i = 1,2,...,C; j = 1,2,...,J_i; k = 1,2,...,K; and m = 1,2,...,M_k. In addition we will denote the number of observations in the jth shot of the ith explosive class by n_{ij} and the total number of observations as N.

For fixed values of i and j, let $\underline{y}_{ij}$ be the vector of n_{ij} observations for the jth shot of the ith explosive class. The model for this vector of observations can be expressed as

$$\underline{y}_{ij} = X_{ij}\underline{\mu}_i + U_{ij}\underline{a}^* + X_{ij}\underline{\beta}^*_{ij} + \underline{e}^*_{ij}, \tag{2-10}$$

where X_{ij} is the matrix of regressor variables (whose rows consist of the $\underline{\phi}'$

row vectors) and $\underline{e}^*_{ij}$ is the vector of ϵ^*_{ijkm} values corresponding to $\underline{y}_{ij}$. Here the vector $\underline{a}^*$ denotes the complete set of α^*_{km} values of interest in the full model, ordered according to gage class. That is

$$\underline{a}^* = (\alpha^*_{11},\ldots,\alpha^*_{1M_1},\alpha^*_{21},\ldots,\alpha^*_{2M_2},\ldots,\alpha^*_{K1},\ldots,\alpha^*_{KM_K})'.$$

The matrix U_{ij}, however, is specific to the jth shot of the ith explosive class and consists of 0's and 1's. Since only a single gage is associated with a particular observation, U_{ij} will have only one 1 per row. An example showing the elements of equation (2-10) explicitly appears in Figure 4.

To assemble the model for the complete set of observations of the ith explosive class we define the observation vector of interest as $\underline{y} = (\underline{y}'_{i1}, \underline{y}'_{i2},\ldots,\underline{y}'_{iJ_i})'$, the corresponding error vector as $\underline{e}^*_i = (\underline{e}^{*\prime}_{i1},\underline{e}^{*\prime}_{i2},\ldots,\underline{e}^{*\prime}_{iJ_i})'$, and the corresponding vector of random performance effects as $\underline{b}^*_i = (\underline{\beta}^{*\prime}_{i1}, \underline{\beta}^{*\prime}_{i2},\ldots,\underline{\beta}^{*\prime}_{iJ_i})'$. Furthermore we define the following matrices: $X_i = (X'_{i1}, X'_{i2},\ldots,X'_{iJ_i})'$, $U_i = (U'_{i1},U'_{i2},\ldots,U'_{iJ_i})'$, and W_i = Block Diag $(X_{i1},X_{i2},\ldots,X_{iJ_i})$. With these the model for $\underline{y}_i$ can be written as

$$\underline{y}_i = X_i\underline{\mu}_i + U_i\underline{a}^* + W_i\underline{b}^*_i + \underline{e}^*_i. \tag{2-11}$$

Finally, we form the full matrix model by defining in analogy with the above $\underline{y} = (\underline{y}'_1,\underline{y}'_2,\ldots,\underline{y}'_C)'$, $\underline{e}^* = (\underline{e}^{*\prime}_1,\underline{e}^{*\prime}_2,\ldots,\underline{e}^{*\prime}_C)'$, $U = (U'_1,U'_2,\ldots,U'_C)'$, and W = Block Diag $(W_1,W_2,\ldots,W_C)$. Also we introduce $\underline{\mu} = (\underline{\mu}'_1,\underline{\mu}'_2,\ldots,\underline{\mu}'_C)'$ and X = Block Diag $(X_1,X_2,\ldots,X_C)$. In terms of these quantities the full matrix model can be represented as

$$
\begin{bmatrix} y_{1111} \\ y_{1113} \\ y_{1122} \\ y_{1127} \\ \vdots \\ y_{1134} \end{bmatrix}
=
\begin{bmatrix} \underline{\phi}'_{1111} \\ \underline{\phi}'_{1113} \\ \underline{\phi}'_{1122} \\ \underline{\phi}'_{1127} \\ \vdots \\ \underline{\phi}'_{1134} \end{bmatrix} \underline{\mu}_1
+
\begin{bmatrix} 1\ 0\ 0\ 0\ \ldots\ 0 \\ 0\ 0\ 1\ 0\ \ldots\ 0 \\ \text{a 1 in } \alpha^*_{22} \text{ column} \\ \text{a 1 in } \alpha^*_{27} \text{ column} \\ \vdots \\ \text{a 1 in } \alpha^*_{34} \text{ column} \end{bmatrix}
\begin{bmatrix} \alpha^*_{11} \\ \vdots \\ \alpha^*_{1M_1} \\ \alpha^*_{21} \\ \vdots \\ \alpha^*_{2M_2} \\ \vdots \\ \alpha^*_{K1} \\ \vdots \\ \alpha^*_{KM_K} \end{bmatrix}
+
\begin{bmatrix} \underline{\phi}'_{1111} \\ \underline{\phi}'_{1113} \\ \underline{\phi}'_{1122} \\ \underline{\phi}'_{1127} \\ \underline{\phi}'_{1134} \end{bmatrix} \underline{\beta}^*_{11}
+
\begin{bmatrix} \epsilon^*_{1111} \\ \epsilon^*_{1113} \\ \epsilon^*_{1122} \\ \epsilon^*_{1127} \\ \epsilon^*_{1134} \end{bmatrix}
$$

$$
\underline{y}_{11} = X_{11}\ \underline{\mu}_1 + U_{11}\ \underline{a}^* + X_{11}\ \underline{\beta}^*_{11} + \underline{e}^*_{11}
$$

Figure 4. Example showing explicit model for observation vector $\underline{y}_{11}$

$$\underline{y} = X\underline{\mu} + U\underline{a}^* + W\underline{b}^* + \underline{e}^*. \qquad (2\text{-}12)$$

As before, an example showing the full matrix model in more explicit form is found in Figure 5.

Equation (2.12) is the matrix counterpart of equation (2.8). The notations employed to describe the various kinds of terms are similar in both cases. In words, (2.12) states that the sample of observations $\underline{y}$ can be viewed as a realization of a random vector in an N dimensional sample space that deviates from the mean point $X\underline{\mu}$ by a vector sum composed of a vector of gage calibration errors $U\underline{a}^*$, a vector of performance variations $W\underline{b}^*$, and a general experimental error vector $\underline{e}^*$. Here, of course, X and W depend upon the transformed reduced distances associated with the sample of observations and U is a matrix of 0's and 1's that links the observations with the random gage effects $\underline{a}^*$.

Viewing $\underline{\mu}$ as a variable, $X\underline{\mu}$ defines a hyperplane in the N dimensional sample space within which the unknown true mean is, by construction of the model, postulated to lie. In Chapter 3 we will be concerned with estimating the position of the true mean on this hyperplane by means of various statistical criteria of choice. For example, the ordinary least squares method selects the orthogonal projection of $\underline{y}$ on the hyperplane as the estimated value of $X\underline{\mu}$. The more precise maximum likelihood and restricted maximum likelihood methods of estimation, that are developed in this report, require explicit representations of the distributions of the random effects in the model. These we now consider.

$$
\begin{bmatrix} \underline{y}_{11} \\ \underline{y}_{12} \\ \vdots \\ \underline{y}_{1J_1} \\ \vdots \\ \underline{y}_{C1} \\ \underline{y}_{C2} \\ \vdots \\ \underline{y}_{CJ_C} \end{bmatrix}
=
\begin{bmatrix} X_{11} & & \\ X_{12} & & 0 \\ \vdots & & \\ X_{1J_1} & & \\ & \ddots & \\ & & X_{C1} \\ 0 & & X_{C2} \\ & & \vdots \\ & & X_{CJ_C} \end{bmatrix}
\begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \\ \vdots \\ \underline{\mu}_C \end{bmatrix}
+
\begin{bmatrix} U_{11} \\ U_{12} \\ \vdots \\ U_{1J_1} \\ \vdots \\ U_{C1} \\ U_{C2} \\ \vdots \\ U_{CJ_C} \end{bmatrix}
\begin{bmatrix} \alpha^*_{11} \\ \vdots \\ \alpha^*_{1M_1} \\ \alpha^*_{21} \\ \vdots \\ \alpha^*_{2M_2} \\ \vdots \\ \alpha^*_{K1} \\ \vdots \\ \alpha^*_{KM_K} \end{bmatrix}
+
\begin{bmatrix} X_{11} & & & & & & & \\ & X_{12} & & & & 0 & & \\ & & \ddots & & & & & \\ & & & X_{1J_1} & & & & \\ & & & & \ddots & & & \\ & & & & & X_{C1} & & \\ & & 0 & & & & X_{C2} & \\ & & & & & & & \ddots & \\ & & & & & & & & X_{CJ_C} \end{bmatrix}
\begin{bmatrix} \underline{\beta}^*_{11} \\ \underline{\beta}^*_{12} \\ \underline{\beta}^*_{1J_1} \\ \underline{\beta}^*_{C1} \\ \underline{\beta}^*_{C2} \\ \underline{\beta}^*_{CJ_C} \end{bmatrix}
+
\begin{bmatrix} \underline{e}^*_{11} \\ \underline{e}^*_{12} \\ \vdots \\ \underline{e}^*_{1J_1} \\ \vdots \\ \underline{e}^*_{C1} \\ \underline{e}^*_{C2} \\ \vdots \\ \underline{e}^*_{CJ_C} \end{bmatrix}
$$

$$\underline{y} = X\,\underline{\mu} + U\,\underline{a}^* + W\,\underline{b}^* + \underline{e}^*$$

Figure 5. Example showing explicit model for complete observation vector $\underline{y}$

2-3 DISTRIBUTIONAL ASSUMPTIONS AND THE DATA VARIANCE COVARIANCE MATRIX

It is reasonable to assume that the sample of observations $\underline{y}$ will have a multivariate normal distribution. Hence, we write

$$\underline{y} \sim N(X\underline{\mu}, V) , \tag{2-13}$$

where $X\underline{\mu}$ is the mean vector of the distribution and V is the N x N variance covariance matrix whose structure we will consider below. As noted earlier the adequacy of this assumption may require a transformation of the measurements. For the usual quantities of interest (peak pressure, decay constant, impulse per unit area, and energy per unit area) experience suggests the logarithm as the appropriate transformation. Use of a logarithmic transformation also supports the assumption of error component additivity as represented in equation (2-3) since calibration effects are essentially multiplicative on the pressure (see Reference 1, p. 183). As in the case of standard regression theory it will be possible to critically examine this and other model assumptions through an analysis of the model residuals.

Implicit in the normality assumption for $\underline{y}$ is the assumption that all random effects of the model are individually normally distributed. Because $\underline{a}^*$, $\underline{b}^*$, and $\underline{e}^*$ are affected by unrelated error sources, they are taken to be mutually independent. We will assume the random effects to have the following multivariate normal distributions,

$$\underline{a}^* \sim N(0, \Gamma\, \sigma^2) \tag{2-14}$$

[1]See footnote 1 on page 1-1.

$$\underline{b}^* \sim N(0, T\,\sigma^2) \tag{2-15}$$

$$\underline{e}^* \sim N(0, I\,\sigma^2) \tag{2-16}$$

where $\Gamma = \text{Diag}\,(\underbrace{\gamma_1,\ldots,\gamma_1}_{M_1},\underbrace{\gamma_2,\ldots,\gamma_2}_{M_2},\ldots,\underbrace{\gamma_K,\ldots,\gamma_K}_{M_K}),$ (2-17)

$$\gamma_k = \sigma^{-2}\,\text{Var}(\alpha^*_{km}),\ m = 1,2,\ldots,M_k, \tag{2-18}$$

and $T = \text{Block Diag}\,(\underbrace{\Theta_1,\ldots,\Theta_1}_{J_1},\underbrace{\Theta_2,\ldots,\Theta_2}_{J_2},\ldots,\underbrace{\Theta_C,\ldots,\Theta_C}_{J_C})$ (2-19)

$$\Theta_i = \sigma^{-2}\,\text{Var}(\underline{\beta}^*_{ij}),\ j = 1,2,\ldots,J_i. \tag{2-20}$$

Hence, Γ is the matrix of gage calibration error variances divided by σ^2 and T is the matrix of the variances and covariances of the random performance variation parameters divided by σ^2. Correlations between the calibration errors of gages with different gage class and calibration indices k and m are assumed to be zero, as are the performance variation parameters of different shots. The use of variance covariance ratios is for later mathematical convenience.

As mentioned earlier, relations (2-17) and (2-18) indicate that the gage calibration error variances are taken to depend on the gage class index k only and not on other factors upon which they might reasonably depend such as reduced range, charge weight, peak pressure, the shock wave decay constant, or individual details of the gages. This is a simplifying assumption that is thought to be reasonably accurate and of practical value.

Using these distributional assumptions and the independence of $\underline{a}^*$, $\underline{b}^*$, and $\underline{e}^*$ we can specify the data variance covariance matrix V. From (2-12) and the rule for forming the variance covariance matrix of a linear function of random variables* we obtain

$$V = H\sigma^2 \tag{2-21}$$

where $$H \equiv I + U\Gamma U' + WTW' . \tag{2-22}$$

It is noted from the block diagonal structures of W and T that WTW' is also block diagonal with blocks $X_{ij}\Theta_i X_{ij}'$ of size $n_{ij} x\ n_{ij}$. The number of observations per shot is usually around 10. The relatively simple structure of H will be exploited below in the task of computing functions of H^{-1}.

*If $\underline{y} = A\underline{x}$ with $E(\underline{x}) = \underline{\mu}$ and $Var(\underline{x}) = \Sigma$, it is easily shown that $E(y) = A\underline{\mu}$ and $Var(\underline{y}) = E[(\underline{y} - E(\underline{y}))(\underline{y} - E(\underline{y}))'] = A\Sigma A'$.

CHAPTER 3

ESTIMATION OF THE MODEL PARAMETERS

We require a general estimation method that will produce efficient estimates of the model parameters, and provide a means for drawing statistical inferences concerning the constants which characterize the larger populations from which the sample is drawn. Of the approaches most suited to mixed models, the maximum likelihood method is the more conceptually and theoretically attractive.* In the past, choice of this method has been often avoided because of the relatively heavy computational burden it involves; but interest has been stimulated in recent years by the development of faster computers, more rapid computational techniques, and various theoretical advances (see References 10 through 16 and the survey paper by Harville, Reference 5). Of the latter, the

*See Searle (Reference 8 and Reference 9, p. 458) for discussions of the suitability of different approach to mixed model estimation.

[8]Searle, S. R., "Topics in Variance Component Estimation," Biometrics, Vol. 27, 1971, p. 1.

[9]Searle, S. R., Linear Models (New York: John Wiley & Sons, Inc., 1971).

[10]Hartley, H. O., and Rao, J. N. K., "Maximum-Likelihood Estimation for the Mixed Analysis of Variance Model," Biometrika, Vol. 54. 1 and 2, 1967, p. 93.

[11]Hemmerle, W. J., and Hartley, H. O., "Computing Maximum Likelihood Estimates for the Mixed A. O. V. Model Using the W Transformation," Technometrics, Vol. 15, No. 4, 1973, p. 819.

[12]Hemmerle, W. J., and Lorens, J. A., "Improved Algorithm for the W-Transform in Variance Component Estimation," Technometrics, Vol. 18, No. 2, 1976, p. 207.

concept of restricted maximum likelihood (REML) estimators will be of interest in the present effort. The REML method, which was extended to the general mixed model by Patterson and Thompson (Reference 17), effectively corrects the ML estimators of variance components for losses in degrees of freedom due to estimation of the fixed effects. In this report we will develop formulae useful for the calculation of both ML and REML estimates. In this effort we will adapt and follow closely the results obtained by Harville (Reference 5).

3-1 MAXIMUM LIKELIHOOD AND OTHER ESTIMATORS OF THE SIMILITUDE EQUATION PARAMETERS $\underline{\mu}$.

Statistical estimators of the fixed effects in linear models are all closely related. Hence, we will use a discussion of the maximum likelihood estimator of $\underline{\mu}$ to motivate a brief discussion of other estimators of the fixed effects. In this manner some of the advantages of ML or REML estimators over the ordinary least squares (OLS) estimators, currently used to obtain estimates of explosive similitude equation parameters, can be highlighted.

[13]Jennrich, R. I., and Sampson, P. F., "Newton-Raphson and Related Algorithms for Maximum Likelihood Variance Component Estimation," Technometrics, Vol. 18, No. 1, 1976, p. 11.

[14]Harville, D. A., "Some Useful Representations for Constrained Mixed-Model Estimation," Journal of the American Statistical Assoc., Vol. 74, No. 365, 1979, p. 200.

[15]Corbeil, R. R., and Searle, S. R., "Restricted Maximum Likelihood (REML) Estimation of Variance Components in the Mixed Model," Technometrics, Vol. 18, No. 1, 1976, p. 31.

[16]Corbeil, R. R., and Searle, S. R., "A Comparison of Variance Component Estimators," Biometrics, Vol. 32, 1976, p. 779.

[5]See footnote 5 on page 1-4.

[17]Patterson, H. D., and Thompson, R., "Recovery of Inter-Block Information When Block Sizes are Unequal," Biometrika, Vol. 58, No. 3, 1971, p. 545.

The likelihood function L is defined as the probability of the observed sample $\underline{y}$ under the assumption that its random behavior is governed by some particular, specified family of distributions. In the present case we have assumed that $\underline{y}$ has a multivariate normal distribution with unknown mean $X\underline{\mu}$ and variance covariance matrix V. Hence,

$$L = (2\pi)^{-N/2} \left|V\right|^{-1/2} \exp[-(\underline{y}-X\mu)'V^{-1}(\underline{y}-X\underline{\mu})/2], \tag{3-1}$$

where $\underline{y}$ denotes the realized sample. L is regarded as a function of the variance and covariance parameters in V (given above in equations (2-17) through (2.20)) and the similitude equation parameters $\underline{\mu}$.

Maximum likelihood estimates are those values of the parameters that maximize L in such a way that L is guaranteed positive not only for the observed values of $\underline{y}$ but for all possible realizations of the sample. In our case this means that the maximization of L is subject to the constraints that $\hat{\sigma}^2 > 0$ and $\hat{\Gamma}$ and $\hat{T}$ be positive definite. In our notation the circumflex (^) will be used to denote the maximum likelihood estimator.

Rather than maximizing L directly one usually works with the log-likelihood function $\lambda = \log L$, which, upon inserting $H\sigma^2$ for V, is found to be

$$\lambda = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log\sigma^2 - \frac{1}{2}\log\left|H\right| - \frac{1}{2\sigma^2}(y-X\underline{\mu})'H^{-1}(\underline{y}-X\underline{\mu}). \tag{3-2}$$

Parameters that maximize λ also maximize L.

For the present we will use $\underline{\theta}$ to denote the unknown variance and covariance ratios in H. The functional dependence of λ may then be summarized as $\lambda(\underline{\mu},\underline{\theta},\sigma^2)$. Following Harville (Reference 5), we will obtain the maximum likelihood estimates $\hat{\underline{\mu}}$, $\hat{\sigma}^2$, and $\hat{\underline{\theta}}$ in three steps:

(1) obtain the function $\overline{\underline{\mu}}(\underline{\theta})$ by maximizing λ with respect to $\underline{\mu}$ for an arbitrary fixed value of θ ($\overline{\underline{\mu}}(\underline{\theta})$ does not involve σ^2),

(2) obtain the ML estimates $\hat{\underline{\theta}}$ and $\hat{\sigma}^2$ by the constrained maximization of $\lambda^*(\underline{\theta},\sigma^2) \equiv \lambda(\overline{\underline{\mu}}(\underline{\theta}),\underline{\theta},\sigma^2)$,

(3) obtain the ML estimate of $\underline{\mu}$ from

$$\hat{\underline{\mu}} = \overline{\underline{\mu}}(\hat{\underline{\theta}}). \tag{3-3}$$

The same approach will be taken to obtain REML estimators $\tilde{\underline{\mu}},\tilde{\underline{\theta}}$, and $\tilde{\sigma}^2$ except for the use of a restricted log-likelihood function $\lambda\#(\underline{\theta},\sigma^2)$.

As λ is continuous and differentiable it can be maximized with respect to $\underline{\mu}$ for fixed values of $\underline{\theta}$ and σ^2 by solving the system of equations $\partial\lambda/\partial\underline{\mu} = 0$. Using well known rules* for differentiation with respect to vectors, we obtain the function $\overline{\underline{\mu}}(\underline{\theta})$ as a solution of the "normal equations"

$$(X'H^{-1}X)\underline{\mu} = X'H^{-1}\underline{y}. \tag{3-4}$$

---

[5]See footnote 5 on page 1-4.

*We use $\partial\underline{z}'\underline{a}/\partial\underline{z} = \underline{a}$ and $\partial\underline{z}'G\underline{z}/\partial\underline{z} = (G + G')\underline{z}$, where $\underline{a}$ and $\underline{z}$ are arbitrary vectors and G is a square matrix.

Since X is the N x Cp matrix of regressor values it is assumed to have full column rank and the matrix $X'H^{-1}X$ is consequently nonsingular. The normal equations then have the unique solution

$$\underline{\bar{\mu}} = (X'H^{-1}X)^{-1}X'H^{-1}\underline{y}. \tag{3-5}$$

Note here that had we included systematic gage bias terms in the model the matrix $X'H^{-1}X$ would have been singular. In that event no unique solution of the full set of similitude equation parameters would have been possible.

The close relationship between least squares estimators of $\underline{\mu}$ and the maximum likelihood estimator (assuming normality as expressed by (3-1)) is seen by observing that the same function that maximizes L for fixed $\underline{\theta}$ and σ^2 also minimizes the weighted sum of squared residuals given by the quadratic form $(\underline{y}-X\underline{\mu})'V^{-1}(\underline{y}-X\underline{\mu})$. It follows from this that equation (3-5) may be used to represent a number of different statistical estimators of $\underline{\mu}$ obtained by least squares and other methods. Some of these are listed in Table 1 and depend upon the nature of H. When H is formed from the ML or REML estimates of $\underline{\theta}$ one obtains, as stated above, the ML or REML estimates of $\underline{\mu}$ respectively. The other estimators listed depend upon H having known or prescribed forms.

At present, similitude equation parameters are estimated by means of the OLS (ordinary least squares) estimator, which does not involve the data variance covariance matrix H at all. It can be shown that all of the other estimators listed in Table 1 are more accurate (in the sense of smaller variances or mean squared errors) than the OLS estimator when correlations exist among the data and $H \neq I$.

TABLE 1 ESTIMATORS OF $\underline{\mu}$ IN THE MODEL $\underline{y} = X\underline{\mu}+\underline{\delta}$, VAR($\underline{y}$) = $H\sigma^2$ HAVING

THE FORM $\underline{\mu} = (X'H^{-1}X)^{-1}X'H^{-1}\underline{y}$

NAME	ESTIMATOR	CONDITIONS
Maximum Likelihood (ML)	$\hat{\underline{\mu}} = (X'\hat{H}^{-1}X)^{-1}X'\hat{H}^{-1}\underline{y}$	Assumes Normality. H evaluated from ML estimates of $\underline{\theta}$.
Restricted Maximum Likelihood (REML)	$\tilde{\underline{\mu}} = (X'\tilde{H}^{-1}X)^{-1}X'\tilde{H}^{-1}\underline{y}$	Assumes Normality. H evaluated from REML estimates of $\underline{\theta}$.
Generalized (Weighted) Least Squares (GLS)	$\tilde{\mu}_{GLS} = (X'H_o^{-1}X)^{-1}X'H_o^{-1}\underline{y}$	$H = H_o$, a matrix of known constants. Minimizes weighted sum of squared residuals $(\underline{y}-X\underline{\mu})'H_o^{-1}(\underline{y}-X\underline{\mu})$.
Best Linear Unbiased Estimator (BLUE)	$\tilde{\mu}_{BLUE} = (X'H_o^{-1}X)^{-1}X'H_o^{-1}\underline{y}$	$H = H_o$, a matrix of known constants. Minimizes the variance of all unbiased estimators of $\underline{\mu}$ linear in $\underline{y}$.
Ordinary Least Squares (OLS)	$\tilde{\mu}_{OLS} = (X'X)^{-1}X'y$	$H = I$, data correlations zero, variances equal and constant. Minimizes weighted sum of squared residuals $(\underline{y}-X\underline{\mu})'(\underline{y}-X\underline{\mu})$.

3-2 THE EXTENDED NORMAL EQUATIONS

In practice it is usually computationally more attractive to find formulations for $\underline{\mu}$ which do not explicitly require the inversion of the N x N variance covariance ratio matrix H (or equivalently V) as does equation (3-5). There are several ways by which this can be achieved that require the inversion of matrices smaller in size. All of these involve about the same computational effort. A particularly useful formulation was published by Henderson et al. (Reference 18) and was investigated extensively by Harville (see References 5, 14, and 19). In addition to giving estimates of the fixed effects $\underline{\mu}$, this formulation also gives estimates of the means of the random effects $\underline{a}^*$ and $\underline{b}^*$ that are conditional upon $\underline{y}$. These can be regarded as estimates of the unknown values, $\underline{a}$ and $\underline{b}$, that are actually realized by $\underline{a}^*$ and $\underline{b}^*$ in the sample. We will find important uses for these in the applications below. Furthermore, Harville (Reference 5) has shown how ML and REML estimates of the variance covariance components $\underline{\theta}$ and σ^2 can be obtained from these results with little additional effort.

In order that we might make direct use of these results, we rewrite the model equation (2-12) as

---

[18]Henderson, C. R., Kempthorne, O., Searle, S. R., and Von Krasigk, C. N., "Estimation of Environmental and Genetic Trends from Records Subject to Culling," Biometrics, Vol. 15, 1959, p. 192.

[5]See footnote 5 on page 1-4.

[14]See footnote 14 on page 3-2.

[19]Harville, D. A., "Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects," The Annals of Statistics, Vol. 4, No. 2, 1976, p. 384.

$$\underline{y} = X\underline{\mu} + Z\underline{v}^* + \underline{e}^*, \tag{3-6}$$

and the variance covariance ratio matrix H given in (2-22) as

$$H = I + ZDZ'. \tag{3-7}$$

Here, quite obviously, we have defined

$$\underline{v}^* \equiv (\underline{a}^{*\prime}, b^{*\prime})', \tag{3-8}$$

$$Z \equiv [U,W], \tag{3-9}$$

and
$$D \equiv \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix}. \tag{3-10}$$

The work of Henderson et al. (from Searle, Reference 9, p. 461) is based upon the joint likelihood of $\underline{y}$ and $\underline{v}^*$, which can be expressed as

$$g(\underline{y},\underline{v}) = g_1(\underline{y}|\underline{v})g_2(\underline{v})$$

$$= c_o \exp[-\frac{1}{2\sigma^2}(\underline{y}-X\underline{\mu}-Z\underline{v})'(\underline{y}-X\underline{\mu}-Z\underline{v})] \exp[-\frac{1}{2\sigma^2}\underline{v}'D^{-1}\underline{v}], \tag{3-11}$$

where c_o is a constant (function of $\underline{\theta}$ and σ^2). Note that although $\underline{y}$ is indiscriminantly used in this report to denote either the random variable or sample value of the observation vector depending on the context, our notation with regard to $\underline{v}^*$ and $\underline{v}$ is more explicit. In (3-11) $\underline{v}$ indicates the unobservable realized value of $\underline{v}^*$. As it is unknown, it may be regarded as a parameter of the model in the same sense as $\underline{\mu}$. For fixed values of θ and σ^2, maximization of (3-11) with respect to $\underline{\mu}$ and $\underline{v}$ leads to the system of equations

[9] See footnote 9 on page 3-1.

$$A \begin{bmatrix} \underline{\mu} \\ \underline{v} \end{bmatrix} = \begin{bmatrix} X'\underline{y} \\ Z'\underline{y} \end{bmatrix} , \qquad (3\text{-}12)$$

where

$$A \equiv \begin{bmatrix} X'X & X'Z \\ Z'X & D^{-1}+Z'Z \end{bmatrix} . \qquad (3\text{-}13)$$

Equations (3-12) are referred to as the "extended normal equations" or also as the "mixed model equations." The solution to (3-12) gives the maximum likelihood estimates of $\underline{\mu}$ and $\underline{v}$ for fixed, arbitrary values of $\underline{\theta}$ and σ^2. It can be shown (e.g., see Serle, Reference 9, pp. 459-462) that the estimator of $\underline{\mu}$ from (3-12) is identical to that given by (3-5) and that the estimator of $\underline{v}$ is equivalently the ML estimator of $E(\underline{v}^*|\underline{y})$. Hence, we can denote solutions of (3-12) as $\underline{\bar{\mu}}$ and $\underline{\bar{v}}$, and as $\underline{\hat{\mu}}$ and $\underline{\hat{v}}$ if the ML estimates of $\underline{\theta}$ and σ^2 have been used.

Solution of (3-12) requires the inversion of the nonsingular coefficient matrix A of size $Cp+\sum_k M_k+p\sum_i J_i$, which equals the number of fixed and random levels in the model. This will be substantially smaller than N. A convenient way of determining A^{-1} is by successively inverting partitioned matrices. We define

$$A^{-1} \equiv \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}'_{12} & \bar{A}_{22} \end{bmatrix} . \qquad (3\text{-}14)$$

[9]See footnote 9 on page 3-1.

Then by substitution of (3-13) and (3-14) into $AA^{-1} = I = A^{-1}A$ (see, e.g., Searle, Reference 20, p. 210 or Westlake, Reference 21, p. 26) one finds

$$\bar{A}_{11} = [X'X-(X'ZB^{-1})Z'X]^{-1} \tag{3-15}$$

$$\bar{A}_{12} = \bar{A}_{11}(X'ZB^{-1}) \tag{3-16}$$

$$\bar{A}_{22} = Q \tag{3-17}$$

where $$B \equiv D^{-1} + Z'Z \tag{3-18}$$

and $$Q \equiv B^{-1} - (X'ZB^{-1})'\bar{A}_{12} \tag{3-19}$$

Since B and Q will be used extensively in the sections below they have been given special notations.

This method of inversion, thus, requires the inversion of B, the inversion of D, and the inversion indicated in (3-15). The inversion of B will be discussed below. The inversion of D is easily obtained since it is composed of a diagonal matrix and a block diagonal matrix consisting of repeated p x p blocks. In practice, p will have a value of 2 or 3 so that the inversion of D is trivial. The matrix $X' - X'ZB^{-1}Z'X$ in (3-15) is symmetric and relatively small in size (Cp x Cp). Hence, it should be possible to obtain $\bar{A}_{11}$ rather

[20]Searle, S. R., Matrix Algebra for the Biological Sciences (New York: John Wiley & Sons, Inc., 1966).

[21]Westlake, J. R., A Handbook of Numerical Matrix Inversion and Solution of Linear Equations (Huntington, NY: Robert E. Krieger Publishing Co., 1975).

rapidly using an approach such as the symmetric Cholesky method (see Westlake, Reference 21, p. 13). Furthermore, the storage of $\bar{A}_{11}$ and $\bar{A}_{22} = Q$ is aided by the symmetry of both matrices.

The inverse of B can also be procured by inverting a partitioned matrix. Using (3-9) and (3-10) we find

$$B = \begin{bmatrix} U'U+\Gamma^{-1} & U'W \\ W'U & W'W+T^{-1} \end{bmatrix} . \tag{3-20}$$

And since B is symmetric we can write its inverse as

$$B^{-1} = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{12} \\ \bar{B}_{12}' & \bar{B}_{22} \end{bmatrix} . \tag{3-21}$$

Now, depending upon the relative values of $\sum_{k=1} M_k$ and $p\sum_{i=1} J_i$ (i.e., the relative sizes of $U'U$ and $W'W$), B^{-1} can be found most efficiently by one of the following methods. If $\sum_k M_k > p\sum_i J_i$, we use

$$\bar{B}_{22} = [W'W+T^{-1}-W'U(U'U+\Gamma^{-1})^{-1}U'W]^{-1} \tag{3-22a}$$

$$\bar{B}_{12} = -(U'U+\Gamma^{-1})^{-1}U'W\bar{B}_{22} \tag{3-23a}$$

$$\bar{B}_{11} = (I-\bar{B}_{12}W'U)(U'U+\Gamma^{-1})^{-1}. \tag{3-24a}$$

But if $\sum_k M_k < p\sum_i J_i$, we use

[21] See footnote 21 on page 3-10.

$$\bar{B}_{11} = [U'U+\Gamma^{-1}-U'W(W'W+T^{-1})^{-1}W'U]^{-1} \tag{3-24b}$$

$$\bar{B}_{12} = -\bar{B}_{11}U'W(W'W+T^{-1})^{-1} \tag{3-23b}$$

$$\bar{B}_{22} = (I-\bar{B}_{12}'U'W)(W'W+T^{-1})^{-1}. \tag{3-22b}$$

One sees that the only difficulties involve equations (3-22a) and (3-24b); the other inverse matrices required are either diagonal (Γ^{-1} and $(U'U+\Gamma^{-1})^{-1}$) or block diagonal with p x p blocks (T^{-1} and $(W'W+T^{-1})^{-1}$). Note also that Γ^{-1} and T^{-1} are readily obtained from (or may be used to obtain) D^{-1}. The inversion indicated in (3-22a) or (3-24b) will be the most time consuming step in the calculation of A^{-1}. Hence, it will probably be important to choose the faster of the two methods indicated above. Since both (3-22a) and (3-24b) are symmetric, the symmetric Cholesky approach is again a good choice of method.

3-3 MAXIMUM LIKELIHOOD ESTIMATORS OF THE VARIANCE COMPONENTS

We obtain the ML estimates of $\underline{\theta}$ and σ^2 by maximizing the function $\lambda^*(\underline{\theta},\sigma^2)$, which from (3-2) is written

$$\lambda^* = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log\sigma^2 - \frac{1}{2}\log|H| - \frac{1}{2\sigma^2}(y-X\underline{\bar{\mu}})'H^{-1}(y-X\underline{\bar{\mu}}). \tag{3-25}$$

Here $\underline{\bar{\mu}}$ is given by (3-5) or obtained from the solution of (3-12). The maximization of (3-25), however, must be carried out subject to constraints that ensure the positive definite property of the estimated variance covariance matrices $\hat{\Gamma}$, $\hat{T}$, and $\hat{\sigma}^2$. Discussion of these constraints is given below in Section 4-2.

Using established rules* for the differentiation of matrices, we can list the derivatives of λ^* with respect to σ^2 and an arbitrary component θ_t of $\underline{\theta}$ as

$$\frac{\partial\lambda^*}{\partial\sigma^2} = \frac{1}{2\sigma^4}(\underline{y}-X\underline{\bar{\mu}})'H^{-1}(\underline{y}-X\underline{\bar{\mu}}) - \frac{N}{2\sigma^2} \tag{3-26}$$

$$\frac{\partial\lambda^*}{\partial\theta_t} = \frac{1}{2\sigma^2}(\underline{y}-X\underline{\bar{\mu}})'H^{-1}\frac{\partial H}{\partial\theta_t}H^{-1}(y-X\underline{\bar{\mu}}) - \frac{1}{2}\,\mathrm{tr}(H^{-1}\frac{\partial H}{\partial\theta_t}). \tag{3-27}$$

Here, (3-27) can be most easily obtained by noting that

$$\partial\lambda^*/\partial\theta_t = (\partial\lambda/\partial\underline{\mu})'(\partial\underline{\mu}/\partial\theta_t) + \partial\lambda/\partial\theta_t = \partial\lambda/\partial\theta_t$$

which is evaluated at $\underline{\mu} = \underline{\bar{\mu}}$. This follows from the requirement that $\partial\lambda/\partial\underline{\mu} = 0$.

The maximum likelihood estimates $\underline{\hat{\theta}}$ and $\hat{\sigma}^2$ satisfy, subject to the constraints, the "likelihood equations" $\partial\lambda^*/\partial\sigma^2 = 0$ and $\partial\lambda^*/\partial\theta_t = 0$ (for all t's), where the derivatives are given by (3-26) and (3-27). Generally, these equations are nonlinear and must be solved iteratively. Recent numerical schemes for computing ML estimates of the parameters in general mixed ANOVA

*We make use of $\partial\log|G|/\partial s = \mathrm{tr}(G^{-1}\partial G/\partial s)$ and $\partial G^{-1}/\partial s = -G^{-1}\partial G/\partial s\, G^{-1}$, where the matrix G depends on the scalar s and its trace tr(G) is the sum of its diagonal elements. Proofs can be obtained from Nering (Reference 22).

[22]Nering, E. D., Linear Algebra and Matrix Theory (New York: John Wiley & Sons, 2nd Ed. 1970).

models have employed the Newton-Raphson (N-R) procedure (see References 11 and 15) or a combination of the N-R procedure and Fisher's method of scoring (Jenrich and Sampson, Reference 13). The method of scoring is a modification of the N-R algorithm which uses the matrix of expected values of the second derivatives in place of the matrix of second derivatives (the Hessian). The N-R method is the more efficient approach and the method of choice when the log-likelihood function λ^* is approximately quadratic, but Jennrich and Sampson suggest that it be backed up by the method of scoring under poor starting conditions or if the N-R iterates begin to diverge. We discuss these techniques more fully in Chapter 4.

Employment of these iterative schemes, thus, requires the availability of second derivatives of λ^* and the expected values of second derivatives. Methods for computing these quantities have been developed by several researchers (References 10, 11, 12, and 14), but the most useful results have been obtained by Harville (Reference 5) who has shown how they may be extracted from the results of the previous section. Hence, the development below is based largely upon Harville's work.

---

[11]See footnote 11 on page 3-1.

[15]See footnote 15 on page 3-2.

[13]See footnote 13 on page 3-2.

[10]See footnote 10 on page 3-1.

[12]See footnote 12 on page 3-1.

[14]See footnote 14 on page 3-2.

[5]See footnote 5 on page 1-4.

The second derivatives of λ^* can be obtained from (3-26) and (3-27) and listed as follows:

$$\frac{\partial^2 \lambda^*}{\partial \sigma^2 \partial \sigma^2} = -\frac{1}{\sigma^6} (\underline{y}-X\underline{\mu})' H^{-1} (\underline{y}-X\underline{\mu}) + \frac{N}{2\sigma^4} \tag{3-28}$$

$$\frac{\partial^2 \lambda^*}{\partial \theta_s \partial \sigma^2} = \frac{1}{2\sigma^4} (\underline{y}-X\underline{\mu})' H^{-1} \left[\frac{\partial H}{\partial \theta_s} (I-2PH)\right] H^{-1} (\underline{y}-X\underline{\mu}) \tag{3-29}$$

$$\frac{\partial^2 \lambda^*}{\partial \theta_s \partial \theta_t} = \frac{1}{2\sigma^2} (\underline{y}-X\underline{\mu})' H^{-1} \left[\frac{\partial^2 H}{\partial \theta_s \partial \theta_t} - 2 \frac{\partial H}{\partial \theta_s} P \frac{\partial H}{\partial \theta_t}\right] H^{-1} (\underline{y}-X\underline{\mu})$$

$$- \frac{1}{2} \operatorname{tr} \left\{ H^{-1} \left[\frac{\partial^2 H}{\partial \theta_s \partial \theta_t} - \frac{\partial H}{\partial \theta_s} H^{-1} \frac{\partial H}{\partial \theta_t}\right]\right\}, \tag{3-30}$$

where $$P \equiv H^{-1} - H^{-1}X(X'H^{-1}X)^{-1}X'H^{-1}. \tag{3-31}$$

In deriving these expressions use is made of the result $\partial P/\partial \theta_s = -P(\partial H/\partial \theta_s)P$, which can be shown by differentiating (3-31) and performing some simple algebra. It should also be noted that $P\underline{y} = H^{-1}(\underline{y}-X\underline{\mu})$ and that P is symmetric.

To obtain the expected values of these derivatives one requires the expression for the mean of a quadratic form (see Searle, Reference 9, p. 55)

$$E(\underline{y}'G\underline{y}) = \operatorname{tr}(Gv) + \underline{\mu}'X'GX\underline{\mu}, \tag{3-32}$$

which is true for a general matrix G. Employing (3-32) and the facts that

[9] See footnote 9 on page 3-1.

$$(PH)^2 = PH, \tag{3-33}$$

$$PX = 0, \tag{3-34}$$

and $$tr(FG) = tr(GF) \tag{3-34}$$

for arbitrary matrices F and G, one finds that

$$E\left(\frac{\partial^2 \gamma^*}{\partial\sigma^2 \partial\sigma^2}\right) = -\frac{1}{\sigma^4}(N-Cp) + \frac{N}{2\sigma^4} \tag{3-36}$$

$$E\left(\frac{\partial^2 \gamma^*}{\partial\theta_s \partial s^2}\right) = -\frac{1}{2\sigma^2}\, tr\left(P\,\frac{\partial H}{\partial\theta_s}\right) \tag{3-37}$$

$$E\left(\frac{\partial^2 \gamma^*}{\partial\theta_s \partial\theta_t}\right) = \frac{1}{2}\, tr\left\{P\left[\frac{\partial^2 H}{\partial\theta_s \partial\theta_t} - 2\,\frac{\partial H}{\partial\theta_s}\, P\, \frac{\partial H}{\partial\theta_t}\right]\right\}$$

$$-\frac{1}{2}\, tr\left\{H^{-1}\left[\frac{\partial^2 H}{\partial\theta_s \partial\theta_t} - \frac{\partial H}{\partial\theta_s}\, H^{-1}\, \frac{\partial H}{\partial\theta_t}\right]\right\}. \tag{3-38}$$

The derivatives in equations (3-26) through (3-30) and the expectations given by (3-36), (3-37), and (3-38) can be simplified by using certain results that we now indicate (see Harville, Reference 5). Many of these results derive from the use of the matrix identity

$$(E+FG)^{-1} \equiv E^{-1}-E^{-1}F(I+GE^{-1}F)^{-1}GE^{-1} \tag{3-39}$$

[5]See footnote 5 on page 1-4.

(obtained by representing $(I+FGE^{-1})^{-1}$ by a geometric series, see Bartlett, Reference 23). Using (3-7) we find

$$Z'H^{-1} = (I+Z'ZD)^{-1}Z' \tag{3-40}$$

$$Z'P = (I+Z'SZD)^{-1}Z'S, \tag{3-41}$$

where $$S \equiv I-X(X'X)^{-1}X' \;. \tag{3-42}$$

From (3-42) it follows that SX = 0. Next we find several expressions for

$$\underline{\bar{\nu}} \equiv D^{-1}\underline{\bar{v}} \;, \tag{3-43}$$

which is a quantity introduced by Harville to facilitate the development of the λ^* derivatives. Using the above relations with the result from (3-12) and (3-14) that $\underline{\bar{v}} = (\bar{A}_{12}'X' + \bar{A}_{22}'Z')\underline{y}$ one can determine that

$$\underline{\bar{\nu}} = (I+Z'ZD)^{-1}Z'(\underline{y}-X\underline{\bar{\mu}}) \tag{3-44}$$

$$= Z'H^{-1}(\underline{y}-X\underline{\bar{\mu}}) \tag{3-45}$$

$$= (I+Z'SZD)^{-1}Z'S\underline{y} \tag{3-46}$$

[23]Bartlett, M. S., "An Inverse Matrix Adjustment Arising in Discriminant Analysis," Annals of Mathematical Statistics, Vol. 22, 1951, p. 107.

Finally, from (3-39) and (3-44) we obtain the useful result

$$H^{-1}(\underline{y}-X\underline{\overline{\mu}}) = \underline{y}-X\underline{\overline{\mu}}-Z\underline{\overline{v}} \equiv \underline{\overline{e}} \quad , \tag{3-47}$$

which is the estimated error vector or vector of residuals, and from (3-5), (3-31), and (3-34) we find

$$(y-X\underline{\overline{\mu}})'H^{-1}(\underline{y}-X\underline{\overline{\mu}}) = \underline{y}'P\underline{y} \quad . \tag{3-48}$$

With these results and remembering that H depends on $\underline{\theta}$ by way of D, we can list the sought after derivatives as

$$\frac{\partial\lambda^*}{\partial\sigma^2} = \frac{1}{\partial\sigma^4}\underline{y}'\underline{\overline{e}} - \frac{N}{\partial\sigma^2} \tag{3-49}$$

$$\frac{\partial\lambda^*}{\partial\theta_t} = \frac{1}{2\sigma^2}\underline{\overline{v}}'\frac{\partial D}{\partial\theta_t}\underline{\overline{v}} - \frac{1}{2}\,\mathrm{tr}\left\{[I+Z'ZD]^{-1}Z'Z\frac{\partial D}{\partial\theta_t}\right\} \tag{3-50}$$

$$\frac{\partial^2\lambda^*}{\partial\sigma^2\partial\sigma^2} = -\frac{1}{\sigma^6}\underline{y}'\underline{\overline{e}} + \frac{N}{2\sigma^4} \tag{3-51}$$

$$\frac{\partial^2\lambda^*}{\partial\theta_t\partial\sigma^2} = -\frac{1}{2}\underline{\overline{v}}'\frac{\partial D}{\partial\theta_t}\underline{\overline{v}} \tag{3-52}$$

$$\frac{\partial^2\lambda^*}{\partial\theta_s\partial\theta_t} = -\frac{1}{\sigma^2}\underline{\overline{v}}'\frac{\partial D}{\partial\theta_s}[I+Z'SZD]^{-1}Z'SZ\frac{\partial D}{\partial\theta_t}\underline{\overline{v}}$$

$$+\frac{1}{2}\,\mathrm{tr}\left\{[I+Z'ZD]^{-1}\,Z'Z\frac{\partial D}{\partial\theta_s}[I+Z'ZD]^{-1}\,Z'Z\frac{\partial D}{\partial\theta_t}\right\} \tag{3-53}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\sigma^2\partial\sigma^2}\right) = -\frac{1}{\sigma^4}(N-Cp) + \frac{N}{2\sigma^4} \tag{3-54}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\theta_t\partial\sigma^2}\right) = -\frac{1}{2\sigma^2}\,\mathrm{tr}\left\{[I+Z'SZD]^{-1}Z'SZ\,\frac{\partial D}{\partial\theta_t}\right\} \tag{3-55}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\theta_s\partial\theta_t}\right) = -\mathrm{tr}\left\{[I+Z'SZD]^{-1}Z'SZ\,\frac{\partial D}{\partial\theta_s}\,[I+Z'SZD]^{-1}Z'SZ\,\frac{\partial D}{\partial\theta_t}\right\}$$

$$+\frac{1}{2}\,\mathrm{tr}\left\{[I+Z'ZD]^{-1}Z'Z\,\frac{\partial D}{\partial\theta_s}\,[I+Z'ZD]\;Z'Z\,\frac{\partial D}{\partial\theta_t}\right\}. \tag{3-56}$$

It remains to express these derivatives in terms of quantities derived in the previous section. Those derivatives that appear complicated depend upon matrices of the forms $[I+Z'SZD]^{-1}Z'SZ\,\frac{\partial D}{\partial\theta_t}$ and $[I+Z'ZD]^{-1}Z'Z\,\frac{\partial D}{\partial\theta_t}$. The latter of these is readily seen to depend on the matrix B^{-1}, a calculational scheme for which was indicated in equations (3-22) through (3-24). Hence, from (3-18) we can write

$$[I+Z'ZD]^{-1}Z'Z\,\frac{\partial D}{\partial\theta_t} = D^{-1}B^{-1}Z'Z\,\frac{\partial D}{\partial\theta_t}\,. \tag{3-57}$$

The former of these matrices has been shown by Harville (Reference 5, p. 326) to depend upon the matrix Q given by equation (3-19). It can be shown by expansion of (3-19) and the use of (3-39) and (3-42) that

$$[I+Z'SZD]^{-1} = D^{-1}Q, \tag{3-58}$$

[5]See footnote 5 on page 1-4.

and, therefore, that

$$[I+Z'SZD]^{-1}Z'SZ\,\frac{\partial D}{\partial\theta_t} = D^{-1}QZ'SZ\,\frac{\partial D}{\partial\theta_t}\quad . \tag{3-59}$$

Now, by means of the identities

$$B'Z'Z \equiv -[B^{-1}D^{-1}-I] \tag{3-60}$$

and $$QZ'SZ \equiv -[QD^{-1}-I]\quad , \tag{3-61}$$

obtained from $I \equiv [D^{-1}+Z'SZ]^{-1}Q^{-1}$ and $I \equiv [D^{-1}+Z'Z]^{-1}B$, it is possible to put (3-57) and (3-59) into very similar forms. We find

$$[I+Z'ZD]^{-1}Z'Z\,\frac{\partial D}{\partial\theta_t} = -[D^{-1}B^{-1}-I]D^{-1}\,\frac{\partial D}{\partial\theta_t} \tag{3-62}$$

and $$[I+Z'SZD]^{-1}Z'SZ\,\frac{\partial D}{\partial\theta_t} = -[D^{-1}Q-I]D^{-1}\,\frac{\partial D}{\partial\theta_t}\quad . \tag{3-63}$$

As a consequence of this similarity of form we now need only continue the development of (3-62), say, and apply the results to (3-63) by substituting Q for B^{-1}.

Equation (3-62) can be simplified by examining the structural details of its matrices and partitioning them according to both gage classes and explosive classes. For D = Block Diag (Γ,T) we define

$$\Gamma_k \equiv \text{Diag}\,\underbrace{(\gamma_k,\ldots,\gamma_k)}_{M_k} = \gamma_k I_{M_k} \tag{3-64}$$

and $\quad T_i \equiv \text{Block Diag}\ (\underbrace{\Theta_i,\ldots,\Theta_i}_{J_i}).$ (3-65)

Hence, from (2-17) and (2-19) the stated partitioning of D may be expressed by

$$\Gamma \equiv \text{Block Diag}\ (\Gamma_1,\Gamma_2,\ldots,\Gamma_K) \tag{3-66}$$

and
$$T \equiv \text{Block Diag}\ (T_1,T_2,\ldots,T_C)\ . \tag{3-67}$$

In a like manner we now partition B^{-1}. From (3-20) and 3-21) we have

$$B^{-1} = \begin{bmatrix} U'U+\Gamma^{-1} & U'W \\ W'U & W'W+T^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} \overline{B}_{11} & \overline{B}_{12} \\ \overline{B}_{21} & \overline{B}_{22} \end{bmatrix} \quad , \tag{3-68}$$

where it is convenient to write $\overline{B}_{21}$ for $\overline{B}_{12}'$. Then the above partitioning may be indicated by writing

$$\overline{B}_{11} = \begin{bmatrix} \overline{B}_{11}(1,1) & \overline{B}_{11}(1,2) & \ldots & \overline{B}_{11}(1,K) \\ \overline{B}_{11}(2,1) & \overline{B}_{11}(2,2) & \ldots & \overline{B}_{11}(2,K) \\ \cdots & \cdots & \cdots & \cdots \\ \overline{B}_{11}(K,1) & \overline{B}_{11}(K,2) & \ldots & \overline{B}_{11}(K,K) \end{bmatrix} , \tag{3-69}$$

$$\overline{B}_{22} = \begin{bmatrix} \overline{B}_{22}(1,1) & \overline{B}_{22}(1,2) & \ldots & \overline{B}_{22}(1,C) \\ \overline{B}_{22}(2,1) & \overline{B}_{22}(2,2) & \ldots & \overline{B}_{22}(2,C) \\ \cdots & \cdots & \cdots & \cdots \\ \overline{B}_{22}(C,1) & \overline{B}_{22}(C,2) & \ldots & \overline{B}_{22}(C,C) \end{bmatrix} , \tag{3-70}$$

$$\overline{B}_{12} = \begin{bmatrix} \overline{B}_{12}(1,1) & \overline{B}_{12}(1,2) & \ldots & \overline{B}_{12}(1,C) \\ \overline{B}_{12}(2,1) & \overline{B}_{12}(2,2) & \ldots & \overline{B}_{12}(2,C) \\ \cdots & \cdots & \cdots & \cdots \\ \overline{B}_{12}(K,1) & \overline{B}_{12}(K,2) & \ldots & \overline{B}_{12}(K,C) \end{bmatrix} , \tag{3-71}$$

with a similar structure for $\overline{B}_{21} = \overline{B}_{12}'$. Note that from the symmetry of B^{-1}, we have $\overline{B}_{11}(k_1,k_2) = (\overline{B}_{11}(k_2,k_1))'$, $\overline{B}_{22}(i_1,i_2) = (\overline{B}_{22}(i_2,i_1))'$, and $\overline{B}_{12}\ (k,i) = (\overline{B}_{21}(i,k))'$.

At this point it is useful to distinguish between the different parameters in $\underline{\theta}$. These were indicated previously in equations (2-18) and (2-20) as the gage class calibration error variance ratios $\gamma_k, k = 1,\ldots,K$ and the explosive class performance variance covariance ratios in $\Theta_i, i = 1,\ldots,C$. Let us now indicate the tth unknown element of Θ_i (and also T_i) as τ_{it}. Here, the actual assignment plan is arbitrary. There will be p! of these parameters for each explosive class.

Choosing $\theta_t = \tau_{it}$, the derivative $\partial D/\partial\theta_t \equiv \partial D/\partial\tau_{it}$ in (3-62) is given by

$$\frac{\partial D}{\partial \tau_{it}} = \begin{bmatrix} 0 & 0 \\ 0 & \partial T/\partial\tau_{it} \end{bmatrix} , \tag{3-72}$$

where $$\frac{\partial T}{\partial \tau_{it}} = \text{Block Diag } (0,\ldots,0, \frac{\partial T_i}{\partial \tau_{it}} ,0,\ldots,0) . \tag{3-73}$$

Consequently,

$$D^{-1} \frac{\partial D}{\partial \tau_{it}} = \text{Block Diag } (0,\ldots,0,T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}} ,0,\ldots,0) . \tag{3-74}$$

Also, the matrix $[D^{-1}B^{-1} - I]$ in (3-62) consists of the rows of the partitioned B^{-1} matrix premultiplied in succession by the diagonal blocks of D^{-1} with identity matrices of the forms I_{M_k} and I_{J_i} subtracted along the diagonal. Postmultiplication of this by (3-74), then, gives a matrix consisting of a single column of submatrices as shown in (3-75). And by a similar argument one obtains (3-76). In these the dashed lines indicate the partitions between

$$
[I + Z'ZD]^{-1}Z'Z\,\frac{\partial D}{\partial \tau_{it}} = -\left[\begin{array}{ccc|cccc}
0 & \cdots & 0 & 0 & \cdots & 0 \quad \Upsilon_1^{-1}\bar{B}_{12}(1,i)\,T_i^{-1}\frac{\partial T_i}{\partial \tau_{it}} & 0 \cdots 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 \quad \Upsilon_2^{-1}\bar{B}_{12}(2,i)\,T_i^{-1}\frac{\partial T_i}{\partial \tau_{it}} & 0 \cdots 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & 0 & \cdots & 0 \quad \Upsilon_K^{-1}\bar{B}_{12}(K,i)\,T_i^{-1}\frac{\partial T_i}{\partial \tau_{it}} & 0 \cdots 0 \\
\hline
0 & \cdots & 0 & 0 & \cdots & 0 \quad T_1^{-1}\bar{B}_{22}(1,i)\,T_i^{-1}\frac{\partial T_i}{\partial \tau_{it}} & 0 \cdots 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & 0 & \cdots & 0 \quad (T_i^{-1}\bar{B}_{22}(i,i) - I_{J_i})T_i^{-1}\frac{\partial T_i}{\partial \tau_{it}} & 0 \cdots 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & 0 & \cdots & 0 \quad T_C^{-1}\bar{B}_{22}(C,i)\,T_i^{-1}\frac{\partial T_i}{\partial \tau_{it}} & 0 \cdots 0
\end{array}\right] \qquad (3\text{-}75)
$$

$$[I+Z'ZD]^{-1}Z'Z\,\frac{\partial D}{\partial \gamma_k} = -\left[\begin{array}{ccccccc|ccc}
0 & \cdots & 0 & \gamma_1^{-1}\bar{B}_{11}(1,k)\,\gamma_k^{-1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdots & 0 & (\gamma_k^{-1}\bar{B}_{11}(k,k) - I_M)\,\gamma_k^{-1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdots & 0 & \gamma_K^{-1}\bar{B}_{11}(K,k)\,\gamma_k^{-1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\hline
0 & \cdots & 0 & T_1^{-1}\bar{B}_{21}(1,k)\,\gamma_k^{-1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & \cdots & 0 & T_2^{-1}\bar{B}_{21}(2,k)\,\gamma_k^{-1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdots & 0 & T_C^{-1}\bar{B}_{21}(C,k)\,\gamma_k^{-1} & 0 & \cdots & 0 & 0 & \cdots & 0
\end{array}\right] \tag{3-76}$$

explosive classes and gage classes. As stated above, these equations apply to equation (3-63) by replacing B^{-1} or $\overline{B}$ matrices by similar Q terms.

These results can now be used to simplify the derivatives in (3-49) through (3-56). In doing so we use the definition

$$\overline{\nu} \equiv \begin{bmatrix} \underline{\overline{\xi}} \\ \underline{\overline{\eta}} \end{bmatrix}, \tag{3-77}$$

where $\underline{\overline{\xi}}$ corresponds to the gage class elements and $\underline{\overline{\eta}}$ to the explosive class elements. Furthermore, we choose $\underline{\overline{\xi}}$ and $\underline{\overline{\eta}}$ to be partitioned in the manner discussed above so that $\underline{\overline{\xi}} \equiv (\underline{\overline{\xi}}_1', \underline{\overline{\xi}}_2', \ldots, \underline{\overline{\xi}}_K')'$ and $\underline{\overline{\eta}} \equiv (\underline{\overline{\eta}}_1', \underline{\overline{\eta}}_2', \ldots, \underline{\overline{\eta}}_C')'$, where $\underline{\overline{\xi}}_k$ and $\underline{\overline{\eta}}_i$ correspond to the kth gage class and the ith explosive class respectively. The derivatives then become

$$\frac{\partial\lambda^*}{\partial\sigma^2} = \frac{1}{2\sigma^4}\,\underline{y}'\underline{\overline{e}} - \frac{N}{2\sigma^2} \tag{3-78}$$

$$\frac{\partial\lambda^*}{\partial\gamma_k} = \frac{1}{2\sigma^2}\,\underline{\overline{\xi}}_k'\underline{\overline{\xi}}_k + \frac{1}{\gamma_k^2}\,\mathrm{tr}\left\{\overline{B}_{11}(k,k)\right\} - \frac{M_k}{\gamma_k} \tag{3-79}$$

$$\frac{\partial\lambda^*}{\partial\tau_{it}} = \frac{-1}{2\sigma^2}\,\underline{\overline{\eta}}_i'\,\frac{\partial T_i}{\partial\tau_{it}}\,\underline{\overline{\eta}}_i + \frac{1}{2}\,\mathrm{tr}\left\{\left[T_i^{-1}\,\overline{B}_{22}(i,i) - I_{J_i}\right]T_i^{-1}\,\frac{\partial T_i}{\partial\tau_{it}}\right\} \tag{3-80}$$

$$\frac{\partial^2\lambda^*}{\partial\sigma^2\,\partial\sigma^2} = -\frac{1}{\sigma^6}\,\underline{y}'\underline{\overline{e}} + \frac{N}{2\sigma^4} \tag{3-81}$$

$$\frac{\partial^2\lambda^*}{\partial\gamma_k\,\partial\sigma^2} = -\frac{1}{2}\,\underline{\overline{\xi}}_k'\underline{\overline{\xi}}_k \tag{3-82}$$

$$\frac{\partial^2 \lambda^*}{\partial \tau_{it} \partial \sigma^2} = -\frac{1}{2} \bar{\eta}_i' \frac{\partial T_i}{\partial \tau_{it}} \bar{\eta}_i \tag{3-83}$$

$$\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \gamma_k} = \frac{1}{\sigma^2 \gamma_k} \bar{\xi}_k' \left[\frac{1}{\gamma_k} Q_{11}(k,k) - I_{M_k} \right] \bar{\xi}_k + \frac{1}{2\gamma_k^2} \operatorname{tr} \left\{ \left[\frac{1}{\gamma_k} \bar{B}_{11}(k,k) - I_{M_k} \right]^2 \right\} \tag{3-84}$$

$$\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \gamma_\ell} = \frac{1}{\sigma^2 \gamma_k \gamma_\ell} \bar{\xi}_k' Q_{11}(k,\ell) \bar{\xi}_\ell + \frac{1}{2\gamma_k^2 \gamma_\gamma^2} \operatorname{tr} \left\{ \bar{B}_{11}(\ell,k) \bar{B}_{11}(k,\ell) \right\} , \quad k \neq \ell \tag{3-85}$$

$$\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \tau_{it}} = \frac{1}{\sigma^2 \gamma_k} \bar{\xi}_k' Q_{12}(k,i) T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}} \bar{\eta}_i + \frac{1}{2\gamma_k^2} \operatorname{tr} \left\{ T_i^{-1} \bar{B}_{21}(i,k) \bar{B}_{12}(k,i) T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}} \right\} \tag{3-86}$$

$$\frac{\partial^2 \lambda^*}{\partial \tau_{is} \partial \tau_{it}} = \frac{1}{\sigma^2} \bar{\eta}_i' \frac{\partial T_i}{\partial \tau_{is}} \left[T_i^{-1} Q_{22}(i,i) - I_{J_i} \right] T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}} \bar{\eta}_i + \frac{1}{2} \operatorname{tr} \left\{ \left[T_i^{-1} \bar{B}_{22}(i,i) - I_{J_i} \right] T_i^{-1} \frac{\partial T_i}{\partial \tau_{is}} \left[T_i^{-1} \bar{B}_{22}(i,i) - I_{J_i} \right] T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}} \right\} \tag{3-87}$$

$$\frac{\partial^2 \lambda^*}{\partial \tau_{hs} \partial \tau_{it}} = \frac{1}{\sigma^2} \bar{n}_h' \frac{\partial T_h}{\partial \tau_{hs}} T_h^{-1} Q_{22}(h,i) T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}} \bar{n}_i$$

$$+ \frac{1}{2} \operatorname{tr} \left\{ T_i^{-1} \bar{B}_{22}(i,h) T_h^{-1} \frac{\partial T_h}{\partial \tau_{hs}} T_h^{-1} \bar{B}_{22}(h,i) T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}} \right\}, \ h \neq i \tag{3-88}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \sigma^2 \partial \sigma^2}\right) = -\frac{1}{\sigma^4}(N-Cp) + \frac{N}{2\sigma^4} \tag{3-89}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \sigma^2}\right) = \frac{1}{2\sigma^2 \gamma_k}\left[\frac{1}{\gamma_k} \operatorname{tr}\left\{Q_{11}(k,k)\right\} - M_k\right] \tag{3-90}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \tau_{it} \partial \sigma^2}\right) = \frac{1}{2\sigma^2} \operatorname{tr}\left\{\left[T_i^{-1} Q_{22}(i,i) - I_{J_i}\right] T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}}\right\} \tag{3-91}$$

$$E\left(\frac{\partial^2 \gamma^*}{\partial \gamma_k \partial \gamma_k}\right) = -\frac{1}{\gamma_k^2} \operatorname{tr}\left\{\left[\frac{1}{\gamma_k} Q_{11}(k,k) - I_{M_k}\right]^2\right\}$$

$$+ \frac{1}{2\gamma_k^2} \operatorname{tr}\left\{\left[\frac{1}{\gamma_k} \bar{B}_{11}(k,k) - I_{M_k}\right]^2\right\} \tag{3-92}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \gamma_\ell}\right) = -\frac{1}{\gamma_k^2 \gamma_\ell^2} \operatorname{tr}\left\{Q_{11}(\ell,k)\, Q_{11}(k,\ell)\right\}$$

$$+ \frac{1}{2\gamma_k^2 \gamma_\ell^2} \operatorname{tr}\left\{\bar{B}_{11}(\ell,k)\, \bar{B}_{11}(k,\ell)\right\}, \ k \neq \ell \tag{3-93}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \tau_{it}}\right) = -\frac{1}{\gamma_k^2} \operatorname{tr}\left\{T_i^{-1} Q_{21}(i,k)\, Q_{12}(k,i)\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}}\right\}$$

$$+ \frac{1}{2\gamma_k^2} \operatorname{tr}\left\{T_i^{-1} \bar{B}_{21}(i,k)\, \bar{B}_{12}(k,i)\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}}\right\} \tag{3-94}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \tau_{is} \partial \tau_{it}}\right) = -\,\mathrm{tr}\left\{(T_i^{-1} Q_{22}(i,i) - I_{J_i})\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{is}}\right.$$

$$\left.(T_i^{-1} Q_{22}(i,i) - I_{J_i})\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}}\right\}$$

$$+ \frac{1}{2}\,\mathrm{tr}\left\{(T_i^{-1} \overline{B}_{22}(i,i) - I_{J_i})\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{is}}\right.$$

$$\left.(T_i^{-1} \overline{B}_{22}(i,i) - I_{J_i})\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}}\right\} \tag{3-95}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \tau_{hs} \partial \tau_{it}}\right) = -\,\mathrm{tr}\left\{T_i^{-1} Q_{22}(i,h)\, T_h^{-1} \frac{\partial T_h}{\partial \tau_{hs}}\right.$$

$$\left.T_h^{-1} Q_{22}(h,i)\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}}\right\}$$

$$+ \frac{1}{2}\,\mathrm{tr}\left\{T_i^{-1} \overline{B}_{22}(i,h)\, T_h^{-1} \frac{\partial T_h}{\partial \tau_{hs}}\right.$$

$$\left.T_h^{-1} \overline{B}_{22}(h,i)\, T_i^{-1} \frac{\partial T_i}{\partial \tau_{it}}\right\},\ h \neq i \tag{3-96}$$

The final simplification of the derivatives is achieved by exploiting the repetition of Θ_i within T_i. As indicated in (2-20) and (3-65) Θ_i is a p x p matrix which is entered along the diagonal of T_i for each of the J_i shots of the explosive class. To take advantage of this structure, then, we further partition $\underline{v}$ and the matrices B^{-1} and Q according to individual shots. We will denote this by subscripting the explosive class index or indices of the particular submatrices. For example, $\overline{B}_{12}(k,i_u)$ will refer to the p columns of $\overline{B}_{12}(k,i)$ that pertain to the uth shot, and $\overline{B}_{22}(h_u,i_v)$ will denote the p x p submatrix of $\overline{B}_{22}(h,i)$ that pertains to the uth shot of the hth explosive class and the vth shot of the ith explosive class.

Employing this notation we can express the derivatives in forms highly suited for computations as follows.

$$\frac{\partial\lambda^*}{\partial\sigma^2} = \frac{1}{2\sigma^4}\,\underline{y}'\underline{\bar{e}} - \frac{N}{2\sigma^2} \tag{3-97}$$

$$\frac{\partial\lambda^*}{\partial\gamma_k} = \frac{1}{2\sigma^2}\,\underline{\bar{\xi}}_k'\underline{\bar{\xi}}_k + \frac{1}{2\gamma_k^2}\,\mathrm{tr}\left\{\bar{B}_{11}(k,k)\right\} - \frac{M_k}{\gamma_k} \tag{3-98}$$

$$\frac{\partial\lambda^*}{\partial\tau_{it}} = \frac{1}{2\sigma^2}\sum_u \underline{\bar{n}}_{i_u}'\,\frac{\partial\Theta_i}{\partial\tau_{it}}\,\underline{\bar{n}}_{i_u}$$

$$+ \frac{1}{2}\sum_u \mathrm{tr}\left\{\left[\Theta_i^{-1}\,\bar{B}_{22}(i_u,i_u) - I_p\right]\Theta_i^{-1}\,\frac{\partial\Theta_i}{\partial\tau_{it}}\right\} \tag{3-99}$$

$$\frac{\partial^2\lambda^*}{\partial\sigma^2\partial\sigma^2} = -\frac{1}{\sigma^6}\,\underline{y}'\underline{\bar{e}} + \frac{N}{2\sigma^4} \tag{3-100}$$

$$\frac{\partial^2\lambda^*}{\partial\gamma_k\partial\sigma^2} = -\frac{1}{2}\,\underline{\bar{\xi}}_k'\underline{\bar{\xi}}_k \tag{3-101}$$

$$\frac{\partial^2\lambda^*}{\partial\tau_{it}\partial\sigma^2} = -\frac{1}{2}\sum_u \underline{\bar{n}}_{i_u}'\,\frac{\partial\Theta_i}{\partial\tau_{it}}\,\underline{\bar{n}}_{i_u} \tag{3-102}$$

$$\frac{\partial^2\lambda^*}{\partial\gamma_k\partial\gamma_k} = \frac{1}{2\sigma^2\gamma_k}\,\underline{\bar{\xi}}_k'\left[\frac{1}{\gamma_k}\,Q_{11}(k,k) - I_{M_k}\right]\underline{\bar{\xi}}_k$$

$$+ \frac{1}{2\gamma_k^2}\,\mathrm{tr}\left\{\left[\frac{1}{\gamma_k}\,\bar{B}_{11}(k,k) - I_{M_k}\right]^2\right\} \tag{3-103}$$

$$\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \gamma_\ell} = \frac{1}{2\sigma^2 \gamma_k \gamma_\ell} \underline{\xi}_k' \, Q_{11}(k,\ell) \, \underline{\xi}_\ell$$

$$+ \frac{1}{2\gamma_k^2 \gamma_\ell^2} \operatorname{tr} \left\{ \overline{B}_{11}(\ell,k) \, \overline{B}_{11}(k,\ell) \right\}, \quad k \neq \ell \tag{-104}$$

$$\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \tau_{it}} = \frac{1}{2\sigma^2 \gamma_k} \sum_u \underline{\xi}_k' \, Q_{12}(k,i_u) \, \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \, \overline{\underline{n}}_{i_u}$$

$$+ \frac{1}{2\gamma_k^2} \sum_u \operatorname{tr} \left\{ \Theta_i^{-1} \, \overline{B}_{21}(i_u,k) \, \overline{B}_{12}(k,i_u) \, \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \right\} \tag{3-105}$$

$$\frac{\partial^2 \lambda^*}{\partial \tau_{is} \partial \tau_{it}} = \frac{1}{2\sigma^2} \sum_u \sum_v \overline{\underline{n}}_{i_u}' \frac{\partial \Theta_i}{\partial \tau_{is}} \left[\Theta_i^{-1} Q_{22}(i_u,i_v) - I_p \right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \, \overline{\underline{n}}_{i_v}$$

$$+ \frac{1}{2} \sum_u \operatorname{tr} \left\{ \left[\Theta_i^{-1} \, \overline{B}_{22}(i_u,i_u) - I_p \right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{is}} \right.$$

$$\left. \left[\Theta_i^{-1} \, \overline{B}_{22}(i_u,i_u) - I_p \right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \right\} \tag{3-106}$$

$$\frac{\partial^2 \lambda^*}{\partial \tau_{hs} \partial \tau_{it}} = \frac{1}{2\sigma^2} \sum_u \sum_v \overline{\underline{n}}_{h_u}' \frac{\partial \Theta_h}{\partial \tau_{hs}} \Theta_h^{-1} Q_{22}(h_u,i_v) \, \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \, \overline{\underline{n}}_{i_v}$$

$$+ \frac{1}{2} \sum_u \operatorname{tr} \left\{ \Theta_i^{-1} \, \overline{B}_{22}(i_u,h_u) \, \Theta_h^{-1} \frac{\partial \Theta_h}{\partial \tau_{hs}} \right.$$

$$\left. \Theta_h^{-1} \, \overline{B}_{22}(h_u,i_u) \, \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \right\}, \quad h \neq i \tag{3-107}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \sigma^2 \partial \sigma^2} \right) = -\frac{1}{\sigma^4}(N - Cp) + \frac{N}{2\sigma^4} \tag{3-108}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial \gamma_k \partial \sigma^2} \right) = \frac{1}{2\sigma^2 \gamma_k} \left[\frac{1}{\gamma_k} \operatorname{tr} \left\{ Q_{11}(k,k) \right\} - M_k \right] \tag{3-109}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\tau_{it}\partial\sigma^2}\right) = \frac{1}{2\sigma^2}\sum_u \mathrm{tr}\left\{\left[\Theta_i^{-1} Q_{22}(i_u,i_u) - I_p\right]\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{it}}\right\} \tag{3-110}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\gamma_k\partial\gamma_k}\right) = -\frac{1}{\gamma_k^2}\mathrm{tr}\left\{\left[\frac{1}{\gamma_k}Q_{11}(k,k) - I_{M_k}\right]^2\right\}$$

$$+\frac{1}{2\gamma_k^2}\mathrm{tr}\left\{\left[\frac{1}{\gamma_k}\overline{B}_{11}(k,k) - I_{M_k}\right]^2\right\} \tag{3-111}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\gamma_k\partial\gamma_\ell}\right) = -\frac{1}{\gamma_k^2\gamma_\ell^2}\mathrm{tr}\left\{Q_{11}(\ell,k)\,Q_{11}(k,\ell)\right\}$$

$$+\frac{1}{2\gamma_k^2\gamma_\ell^2}\mathrm{tr}\left\{\overline{B}_{11}(\ell,k)\,\overline{B}_{11}(k,\ell)\right\}\quad,\quad k\neq\ell \tag{3-112}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\gamma_k\partial\tau_{it}}\right) = -\frac{1}{\gamma_k^2}\sum_u \mathrm{tr}\left\{\Theta_i^{-1} Q_{21}(i_u,k)\,Q_{12}(k,i_u)\,\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{it}}\right\}$$

$$+\frac{1}{2\gamma_k^2}\sum_u \mathrm{tr}\left\{\Theta_i^{-1}\overline{B}_{21}(i_u,k)\,\overline{B}_{12}(k,i_u)\,\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{it}}\right\} \tag{3-113}$$

$$E\left(\frac{\partial^2\lambda^*}{\partial\tau_{is}\partial\tau_{it}}\right) = -\sum_u \mathrm{tr}\left\{\left[\Theta_i^{-1} Q_{22}(i_u,i_u) - I_p\right]\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{is}}\right.$$

$$\left.\left[\Theta_i^{-1} Q_{22}(i_u,i_u) - I_p\right]\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{it}}\right\}$$

$$+\frac{1}{2}\sum_u \mathrm{tr}\left\{\left[\Theta_i^{-1}\overline{B}_{22}(i_u,i_u) - I_p\right]\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{is}}\right.$$

$$\left.\left[\Theta_i^{-1}\overline{B}_{22}(i_u,i_u) - I_p\right]\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{it}}\right\} \tag{3-114}$$

$$E\left(\frac{\partial^2 \lambda^*}{\partial\tau_{hs}\partial\tau_{it}}\right) = -\sum_u \mathrm{tr}\left\{\Theta_i^{-1} Q_{22}(i_u,h_u)\right.$$

$$\left.\Theta_h^{-1}\frac{\partial\Theta_h}{\partial\tau_{hs}}\Theta_h^{-1} Q_{22}(h_u,i_u)\,\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{it}}\right\}$$

$$+\frac{1}{2}\sum_u \mathrm{tr}\left\{\Theta_i^{-1}\bar{B}_{22}(i_u,h_u)\right.$$

$$\left.\Theta_h^{-1}\frac{\partial\Theta_h}{\partial\tau_{hs}}\Theta_h^{-1}\bar{B}_{22}(h_u,i_u)\,\Theta_i^{-1}\frac{\partial\Theta_i}{\partial\tau_{it}}\right\}, \quad h \neq i \qquad (3\text{-}115)$$

In these expressions the sums on u and v range from 1 to the number of shots in the explosive class (J_i for the ith class).

3-4 RESTRICTED MAXIMUM LIKELIHOOD ESTIMATORS OF THE VARIANCE COMPONENTS

It is well known that maximum likelihood estimators of variance components, symbolized above by $\hat{\sigma}^2$ and $\hat{\underline{\theta}}$, are biased, i.e., $E(\hat{\underline{\sigma}}^2) \neq \sigma^2$ and $E(\hat{\underline{\theta}}) \neq \underline{\theta}$. Hence, the distributions of the estimators are not centered, in the sense of the means, about the true values of the quantities being estimated. For small data samples the bias can lead to substantial errors in estimation. Corbeil and Searle (Reference 15), for example, show for a sample of size 16 that ML estimators can underpredict estimators that are known to be unbiased by a factor of two. The restricted maximum likelihood (REML) method is an attempt to overcome this problem of bias. It has been demonstrated to produce unbiased estimates when applied to a number of different linear models (see Harville, Reference 5), and the property of unbiasedness is believed by some to be a

[15]See footnote 15 on page 3-2.

[5]See footnote 5 on page 1-4.

general property of the method. Corbeil and Searle (Reference 16), however, have shown that this may come at the expense of estimator efficiency. Because of the unbiasedness property of the REML variance components estimators and their close association with the ML estimators, we include them in this report.

The REML estimators of σ^2 and $\underline{\theta}$ are based upon the likelihood function associated with certain linearly independent combinations of the data which possess zero expectations. Such sums of the observations, of which there is a total of N−Cp in the present case, are known as the error contrasts. It is argued (Patterson and Thompson, Reference 17) that they may be thought of as containing all of the variance component information and should, therefore, form the basis of estimation.

If we let $R\underline{y}$ denote a particular set of the error contrasts, it has been indicated by Harville* (Reference 5) that to within an additive constant the log-likelihood function for $R\underline{y}$ may be written as

$$\lambda^{\#} = -\frac{(N-Cp)}{2}\log\sigma^2 - \frac{1}{2}\log|H| - \frac{1}{2}\log|X'H^{-1}X| - \frac{1}{2\sigma^2}(\underline{y}-X\underline{\mu})'H^{-1}(\underline{y}-X\underline{\mu}). \quad (3\text{-}116)$$

---

[16]See footnote 16 on page 3-2.

[17]See footnote 17 on page 3-2.

*Harville cites his 1974 paper (Reference 24) for this derivation, but it is derived in a Bayesian context. However, the author has been able to verify this expression in a classical setting using rather straightforward variable transformation theory along with several matrix relations published by Harville in his 1974 paper.

[24]Harville, D. A., "Bayesian Inference for Variance Components Using Only Error Contrasts," Biometrika, Vol. 61, 1974, p. 38.

[5]See footnote 5 on page 1-4.

The REML estimates of $\underline{\theta}$ and σ^2 are those values that maximize (3-116). We now show how the derivatives required for the maximization of $\lambda^{\#}$ by the N-R and Fisher scoring optimization schemes may be obtained from the derivatives of λ^{*}.

In the same manner that we obtained the λ^{*} derivatives given by (3-26) through (3-30), and (3-36) through (3-38) we can write

$$\frac{\partial\lambda^{\#}}{\partial\sigma^2} = \frac{1}{2\sigma^4}\,(\underline{y}-X\underline{\bar{\mu}})'H^{-1}(\underline{y}-X\underline{\bar{\mu}}) - \frac{1}{2\sigma^2}\,(N-Cp) \tag{3-117}$$

$$\frac{\partial\lambda^{\#}}{\partial\theta_t} = \frac{1}{2\sigma^2}\,(\underline{y}-X\underline{\bar{\mu}})'H^{-1}\,\frac{\partial H}{\partial\theta_t}\,H^{-1}(y-X\underline{\bar{\mu}}) - \frac{1}{2}\,\mathrm{tr}\,\left(P\,\frac{\partial H}{\partial\theta_t}\right) \tag{3-118}^{*}$$

$$\frac{\partial^2\lambda^{\#}}{\partial\sigma^2\,\partial\sigma^2} = -\,\frac{1}{\sigma^6}\,(\underline{y}-X\underline{\bar{\mu}})'H^{-1}(\underline{y}-X\underline{\bar{\mu}}) + \frac{1}{2\sigma^4}\,(N-Cp) \tag{3-119}$$

$$\frac{\partial^2\lambda^{\#}}{\partial\theta_s\,\partial\sigma^2} = \frac{1}{2\sigma^4}\,(y-X\underline{\bar{\mu}})'H^{-1}\left[\frac{\partial H}{\partial\theta_s}\,(I-2PH)\right]\,H^{-1}(\underline{y}-X\underline{\bar{\mu}}) \tag{3-120}$$

$$\frac{\partial^2\lambda^{\#}}{\partial\theta_s\,\partial\theta_t} = \frac{1}{2\sigma^2}\,(\underline{y}-X\underline{\bar{\mu}})'H^{-1}\left[\frac{\partial^2 H}{\partial\theta_s\,\partial\theta_t} - 2\,\frac{\partial H}{\partial\theta_s}\,P\,\frac{\partial H}{\partial\theta_t}\right]H^{-1}(\underline{y}-X\underline{\bar{\mu}})$$

$$-\,\frac{1}{2}\,\mathrm{tr}\left\{P\left[\frac{\partial^2 H}{\partial\theta_s\,\partial\theta_t} - \frac{\partial H}{\partial\theta_s}\,P\,\frac{\partial H}{\partial\theta_t}\right]\right\} \tag{3-121}$$

$$E\left(\frac{\partial^2\lambda^{\#}}{\partial\sigma^2\,\partial\sigma^2}\right) = -\,\frac{1}{\sigma^4}\,(N-Cp) + \frac{(N-Cp)}{2\sigma^4} = -\,\frac{1}{2\sigma^4}\,(N-Cp) \tag{3-122}$$

---

*Note that if $\lambda^{\#} \equiv g^{\#}(\sigma^2,\ \underline{\theta},\ \underline{\bar{\mu}}(\underline{\theta}))$ and $g = g^{\#}(\sigma^2,\ \underline{\theta},\ \underline{\mu})$, that $\underline{\bar{\mu}}$ is a solution to $\partial g/\partial\underline{\mu} = 0$. Hence, $\partial\lambda^{\#}/\partial\theta_t = \partial g/\partial\theta_t$ evaluated at $\underline{\mu} = \underline{\bar{\mu}}$.

$$E\left(\frac{\partial^2\lambda^{\#}}{\partial\theta_s\partial\sigma^2}\right) = -\frac{1}{2\sigma^2}\,\mathrm{tr}\left\{P\,\frac{\partial H}{\partial\theta_s}\right\} \tag{3-123}$$

$$E\left(\frac{\partial^2\lambda^{\#}}{\partial\theta_s\partial\theta_t}\right) = \frac{1}{2}\,\mathrm{tr}\left\{P\left[\frac{\partial^2 H}{\partial\theta_s\partial\theta_t} - 2\,\frac{\partial H}{\partial\theta_s}\,P\,\frac{\partial H}{\partial\theta_t}\right]\right\}$$

$$-\frac{1}{2}\,\mathrm{tr}\left\{P\left[\frac{\partial^2 H}{\partial\theta_s\partial\theta_t} - \frac{\partial H}{\partial\theta_s}\,H^{-1}\,\frac{\partial H}{\partial\theta_t}\right]\right\} = -\frac{1}{2}\,\mathrm{tr}\left\{P\,\frac{\partial H}{\partial\theta_s}\,P\,\frac{\partial H}{\partial\theta_t}\right\} \tag{3-124}$$

It is observed that equations (3-117) through (3-124) differ only slightly from equations (3-26) through (3-30) and (3-36) through (3-38). The expectations of derivatives are shown above both in a way that emphasizes this similarity and also in their algebraically simplified forms. Close inspection of these results reveals that we can immediately write down the REML equivalents to equations (3-97) through (3-115) by simply using the submatrices of Q where those of B^{-1} appear and by employing (N-Cp) in certain places where only N appears. By proceeding in this manner and simplifying the results we obtain the computational forms of the derivatives as

$$\frac{\partial\lambda^{\#}}{\partial\sigma^2} = \frac{1}{2\sigma^4}\,\underline{y}'\underline{\bar{e}} - \frac{1}{2\sigma^2}\,(N-Cp) \tag{3-125}$$

$$\frac{\partial\lambda^{\#}}{\partial\gamma_k} = \frac{1}{2\sigma^2}\,\underline{\bar{\xi}}_k'\underline{\bar{\xi}}_k + \frac{1}{2\gamma_k^2}\,\mathrm{tr}\left\{Q_{11}(k,k)\right\} - \frac{M_k}{2\gamma_k} \tag{3-126}$$

$$\frac{\partial\lambda^{\#}}{\partial\tau_{it}} = \frac{1}{2\sigma^2}\sum_u \underline{\bar{\eta}}_{i_u}'\,\frac{\partial\Theta_i}{\partial\tau_{it}}\,\underline{\bar{\eta}}_{i_u}$$

$$+\frac{1}{2}\sum_u \mathrm{tr}\left\{\left[\Theta_i^{-1}\,Q_{22}(i_u,i_u) - I_p\right]\Theta_i^{-1}\,\frac{\partial\Theta_i}{\partial\tau_{it}}\right\} \tag{3-127}$$

$$\frac{\partial^2 \lambda^\#}{\partial\sigma^2 \partial\sigma^2} = -\frac{1}{\sigma^6} \underline{y}' \underline{\underline{e}} + \frac{1}{2\sigma^4} (N-Cp) \tag{3-128}$$

$$\frac{\partial^2 \lambda^\#}{\partial\gamma_k \partial\sigma^2} = -\frac{1}{2} \underline{\bar{\xi}}_k' \underline{\bar{\xi}}_k \tag{3-129}$$

$$\frac{\partial^2 \lambda^\#}{\partial\tau_{it} \partial\sigma^2} = -\frac{1}{2} \sum_u \bar{\underline{\eta}}_{i_u} \frac{\partial \Theta_i}{\partial\tau_{it}} \bar{\underline{\eta}}_{i_u} \tag{3-130}$$

$$\frac{\partial^2 \lambda^\#}{\partial\gamma_k \partial\gamma_k} = \frac{1}{2\sigma^2 \gamma_k} \bar{\xi}_k' \left[\frac{1}{\gamma_k} Q_{11}(k,k) - I_{M_k}\right] \bar{\xi}_k + \frac{1}{2\gamma_k^2} \operatorname{tr} \left\{\left[\frac{1}{\gamma_k} Q_{11}(k,k) - I_{M_k}\right]^2\right\} \tag{3-131}$$

$$\frac{\partial^2 \lambda^\#}{\partial\gamma_k \partial\gamma_\ell} = \frac{1}{2\sigma^2 \gamma_k \gamma_\ell} \bar{\xi}_k' Q_{11}(k,\ell) \bar{\xi}_\ell + \frac{1}{2\gamma_k^2\gamma_\ell^2} \operatorname{tr} \left\{Q_{11}(\ell,k) Q_{11}(k,\ell)\right\} , \quad k \neq \ell \tag{3-132}$$

$$\frac{\partial^2 \lambda^\#}{\partial\gamma_k \partial\tau_{it}} = \frac{1}{2\sigma^2 \gamma_k} \sum_u \bar{\xi}_k' Q_{12}(k,i_u) \Theta_i^{-1} \frac{\partial \Theta_i}{\partial\tau_{it}} \bar{\eta}_{i_u} + \frac{1}{2\gamma_k^2} \sum_u \operatorname{tr} \left\{\Theta_i^{-1} Q_{21}(i_u,k) Q_{12}(k,i_u) \Theta_i^{-1} \frac{\partial \Theta_i}{\partial\tau_{it}}\right\} \tag{3-133}$$

$$\frac{\partial^2 \lambda^\#}{\partial\tau_{is} \partial\tau_{it}} = \frac{1}{2\sigma^2} \sum_u \sum_v \bar{\eta}_{i_u}' \frac{\partial \Theta_i}{\partial\tau_{is}} \left[\Theta_i^{-1} Q_{22}(i_u,i_v) - I_p\right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial\tau_{it}} \bar{\eta}_{i_v} + \frac{1}{2} \sum_u \operatorname{tr} \left\{\left[\Theta_i^{-1} Q_{22}(i_u,i_u) - I_p\right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial\tau_{is}} \left[\Theta_i^{-1} Q_{22}(i_u,i_u) - I_p\right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial\tau_{it}}\right\} \tag{3-134}$$

$$\frac{\partial^2 \lambda^{\#}}{\partial \tau_{hs} \partial \tau_{it}} = \frac{1}{2\sigma^2} \sum_u \sum_v \bar{n}'_{h_u} \frac{\partial \Theta_h}{\partial \tau_{hs}} \Theta_h^{-1} Q_{22}(h_u, i_v) \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \bar{n}_{i_v}$$

$$+ \frac{1}{2} \sum_u \text{tr} \left\{ \Theta_i^{-1} Q_{22}(i_u, h_u) \Theta_h^{-1} \frac{\partial \Theta_h}{\partial \tau_{hs}} \right.$$

$$\left. \Theta_h^{-1} Q_{22}(h_u, i_u) \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}} \right\} \quad , \quad h \neq i \tag{3-135}$$

$$E\left(\frac{\partial^2 \lambda^{\#}}{\partial \sigma^2 \partial \sigma^2}\right) = - \frac{1}{2\sigma^4} (N-Cp) \tag{3-136}$$

$$E\left(\frac{\partial^2 \lambda^{\#}}{2\gamma_k \partial \sigma^2}\right) = \frac{1}{2\sigma^2 \gamma_k} \left[\frac{1}{\gamma_k} \text{tr} \left\{Q_{11}(k,k)\right\} - M_k\right] \tag{3-137}$$

$$E\left(\frac{\partial^2 \lambda^{\#}}{\partial \tau_{it} \partial \sigma^2}\right) = \frac{1}{2\sigma^2} \sum_u \text{tr} \left\{\left[\Theta_i^{-1} Q_{22}(i_u, i_u) - I_p\right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}}\right\} \tag{3-138}$$

$$E\left(\frac{\partial^2 \lambda^{\#}}{\partial \gamma_k \partial \gamma_k}\right) = - \frac{1}{2\gamma_k^2} \text{tr} \left\{\left[\frac{1}{\gamma_k} Q_{11}(k,k) - I_{M_k}\right]^2\right\} \tag{3-139}$$

$$E\left(\frac{\partial^2 \lambda^{\#}}{\partial \gamma_k \partial \gamma_\ell}\right) = - \frac{1}{2\gamma_k^2 \gamma_\ell^2} \text{tr} \left\{Q_{11}(\ell,k) \, Q_{11}(k,\ell)\right\} \quad , \quad k \neq \ell \tag{3-140}$$

$$E\left(\frac{\partial^2 \lambda^{\#}}{\partial \gamma_k \partial \tau_{it}}\right) = - \frac{1}{2\gamma_k^2} \sum_u \text{tr} \left\{\Theta_i^{-1} Q_{21}(i_u, k) \, Q_{12}(k, i_u) \, \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}}\right\} \tag{3-141}$$

$$E\left(\frac{\partial^2 \lambda^{\#}}{\partial \tau_{is} \partial \tau_{it}}\right) = - \frac{1}{2} \sum_u \text{tr} \left\{\left[\Theta_i^{-1} Q_{22}(i_u, i_u) - I_p\right]\right.$$

$$\left. \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{is}} \left[\Theta_i^{-1} Q_{22}(i_u, i_u) - I_p\right] \Theta_i^{-1} \frac{\partial \Theta_i}{\partial \tau_{it}}\right\} \tag{3-142}$$

$$E\left(\frac{\partial^2\lambda^{\#}}{\partial\tau_{hs}\partial\tau_{it}}\right) = -\frac{1}{2}\sum_{u} \mathrm{tr}\left\{\Theta_i^{-1}\, Q_{22}(i_u, h_u)\, \Theta_h^{-1}\, \frac{\partial\Theta_h}{\partial\tau_{hs}}\, \Theta_h^{-1}\, Q_{22}(h_u, i_u)\, \Theta_i^{-1}\, \frac{\partial\Theta_i}{\partial\tau_{it}}\right\},$$

$$h \neq i \qquad (3\text{-}143)$$

As before, the sums on u and v range from 1 to the number of shots in the explosive class.

CHAPTER 4

NUMERICAL METHODS

Following Jennrich and Sampson (Reference 13) we propose the use of the combined Newton-Raphson and Fisher scoring procedures to obtain the ML or REML estimates of the variance compo ents. Satisfaction of the constraints on the iterated variance covariance matrices will be handled by the interior penalty function technique. While the expressions below may be used interchangeably to obtain either ML or REML estimates, they will, for the sake of brevity, be given in most cases in terms of the ML notation only.

4-1 UNCONSTRAINED OPTIMIZATION

Concise descriptions of the Newton-Raphson and Fisher scoring algorithms may be stated in terms of the parameter vector

$$\underline{\omega} \equiv (\sigma^2, \gamma_1, \ldots, \gamma_k, \tau_{11}, \ldots, \tau_{1\pi}, \ldots, \tau_{C1}, \ldots, \tau_{C\pi})' \quad , \tag{4-1}$$

where $\pi \equiv p!$. The N-R method is an iterative procedure that corrects an initial guess $\underline{\omega}_o$, or the value of the iterate after i steps, $\underline{\omega}_i$, by an amount

$$\underline{\Delta\omega_i} = -\mathcal{H}_i^{-1} \underline{\nabla\lambda_i}^* \quad , \tag{4-2}$$

[13]See footnote 13 on page 3-2.

where $\underline{\Delta\omega}_i = \underline{\omega}_{i+1} - \underline{\omega}_i$, $\underline{\nabla\lambda}^*$ is the gradient vector

$$\nabla\lambda^* \equiv \left(\frac{\partial\lambda^*}{2\sigma^2}, \frac{\partial\lambda^*}{\partial\gamma 1}, \ldots, \frac{\partial\lambda^*}{\partial\gamma K}, \frac{\partial\lambda^*}{\partial\tau 11}, \ldots, \frac{\partial\lambda^*}{\partial\tau 1\pi}, \ldots, \frac{\partial\lambda^*}{\partial\tau C1}, \ldots, \frac{\partial\lambda^*}{\partial\tau C\pi}\right)', \tag{4-3}$$

and $\mathcal{H}$ is the Hessian matrix defined by equation (4-4) shown on the opposite page. Note that in (4-2) $\underline{\nabla\lambda}^*$ and $\mathcal{H}$ are evaluated at $\underline{\omega}_i$. The method of scoring is identical to the N-R algorithm except that $E(\mathcal{H})$ is used in place of $\mathcal{H}$.

Both the N-R procedure and the method of scoring approximate the λ^* function at each step of the iteration by a quadratic function. In the case of the N-R method this is

$$\lambda^*(\underline{\omega}_{i+1}) = \lambda^*(\underline{\omega}_i) + (\underline{\nabla\lambda}^*_i)' \,\underline{\Delta\omega}_i + \frac{1}{2}(\underline{\Delta\omega}_i)' \mathcal{H}_i \underline{\Delta\omega}_i, \tag{4-5}$$

which is recognized as the Taylor expansion of λ^* about the point $\underline{\omega}_i$ up to the second order term. Upon taking the gradient of $\lambda^*(\underline{\omega}_{i+1})$ we get*

$$\underline{\nabla\lambda}^*(\underline{\omega}_{i+1}) = \underline{\nabla\lambda}^*_i + \mathcal{H}_i \underline{\Delta\omega}_i \quad , \tag{4-6}$$

which must be zero at a maximum. Hence, equating (4-6) to zero and solving for $\underline{\Delta\omega}_i$ gives (4-2). In order for (4-5) to approximate λ^* near a maximum, as opposed to a minimum or saddle point, the Hessian $\mathcal{H}_i$ must be negative definite, i.e., $z'\mathcal{H}z < 0$ for any vector $z \neq 0$ (in particular $\underline{\Delta\omega}_i$), since any excursion away from the maximum must result in a decrease in λ^*.

*Use footnote on page 3-4 and symmetry of $\mathcal{H}$.

$$
\mathcal{H} = \left[\begin{array}{ccccccccccc}
\frac{\partial^2\lambda^*}{\partial\sigma^2\partial\sigma^2} & \frac{\partial^2\lambda^*}{\partial\sigma^2\partial\gamma_1} & \cdots & \frac{\partial^2\lambda^*}{\partial\sigma^2\partial\gamma_K} & \frac{\partial^2\lambda^*}{\partial\sigma^2\partial\tau_{11}} & \cdots & \frac{\partial^2\lambda^*}{\partial\sigma^2\partial\tau_{1\pi}} & \cdots & \frac{\partial^2\lambda^*}{\partial\sigma^2 2\tau_{C1}} & \cdots & \frac{\partial^2\lambda^*}{\partial\sigma^2\partial\tau_{C\pi}} \\
 & \frac{\partial^2\lambda^*}{\partial\gamma_1\partial\gamma_1} & \cdots & \frac{\partial^2\lambda^*}{\partial\gamma_1\partial\gamma_K} & \frac{\partial^2\lambda^*}{\partial\gamma_1\partial\tau_{11}} & \cdots & \frac{\partial^2\lambda^*}{\partial\gamma_1\partial\tau_{1\pi}} & \cdots & \frac{\partial^2\lambda^*}{\partial\gamma_1\partial\tau_{C1}} & \cdots & \frac{\partial^2\lambda^*}{\partial\gamma_1\partial\tau_{C\pi}} \\
 & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & \frac{\partial^2\lambda^*}{\partial\gamma_K\partial\gamma_K} & \frac{\partial^2\lambda^*}{\partial\gamma_K\partial\tau_{11}} & \cdots & \frac{\partial^2\lambda^*}{\partial\gamma_K\partial\tau_{1\pi}} & \cdots & \frac{\partial^2\lambda^*}{\partial\gamma_K\partial\tau_{C1}} & \cdots & \frac{\partial^2\lambda^*}{\partial\gamma_K\partial\tau_{C\pi}} \\
 & & & & \frac{\partial^2\lambda^*}{\partial\tau_{11}\partial\tau_{11}} & \cdots & \frac{\partial^2\lambda^*}{\partial\tau_{11}\partial\tau_{1\pi}} & \cdots & \frac{\partial^2\lambda^*}{\partial\tau_{11}\partial\tau_{C1}} & \cdots & \frac{\partial^2\lambda^*}{\partial\tau_{11}\partial\tau_{C\pi}} \\
 & & & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & & & & \frac{\partial^2\lambda^*}{\partial\tau_{1\pi}\partial\tau_{1\pi}} & \cdots & \frac{\partial^2\lambda^*}{\partial\tau_{1\pi}\partial\tau_{C1}} & \cdots & \frac{\partial^2\lambda^*}{\partial\tau_{1\pi}\partial\tau_{C\pi}} \\
 & & & & & & & \cdots & \cdots & \cdots & \cdots \\
 & & \text{Symmetric} & & & & & & \frac{\partial^2\lambda^*}{\partial\tau_{C1}\partial\tau_{C1}} & \cdots & \frac{\partial^2\lambda^*}{\partial\tau_{C1}\partial\tau_{C\pi}} \\
 & & & & & & & & & \cdots & \cdots \\
 & & & & & & & & & & \frac{\partial^2\lambda^*}{\partial\tau_{C\pi}\partial\tau_{C\pi}}
\end{array}\right] \qquad (4\text{-}4)
$$

The fact that away from the maximum $\mathcal{H}$ need not be negative definite can cause the N-R interate to diverge from the maximum if the starting conditions are poor. It is for this reason that Jennrich and Sampson (Reference 13) employ the method of scoring in the initial step of the iteration and whenever the process appears to diverge. Since $E(\mathcal{H})$ is nonpositive definite (≤ 0) and in our case will most likely be negative definite (<0), the scoring step $\underline{\Delta\omega}_i = -[E(\mathcal{H})]^{-1}\underline{\nabla\lambda}_i^*$ will at least locally always be in the direction of increasing λ^*. To show this we examine the component of $\underline{\Delta\omega}_i$ in the direction of $\underline{\nabla\lambda}_i^*$. Using Figure 6, this is

$$|\underline{\Delta\omega}_i|\cos\phi = (\underline{\Delta\omega}_i)'\underline{\nabla\lambda}_i^*/|\underline{\Delta\lambda}_i^*| = -(\underline{\nabla\lambda}_i^*)'[E(\mathcal{H})]^{-1}\ \underline{\nabla\lambda}_i^*/|\underline{\nabla\lambda}_i^*| \geq 0. \tag{4-7}$$

Hence, $\cos\phi > 0$ and $\underline{\Delta\omega}_i$ has a component in the direction of increasing λ^*.

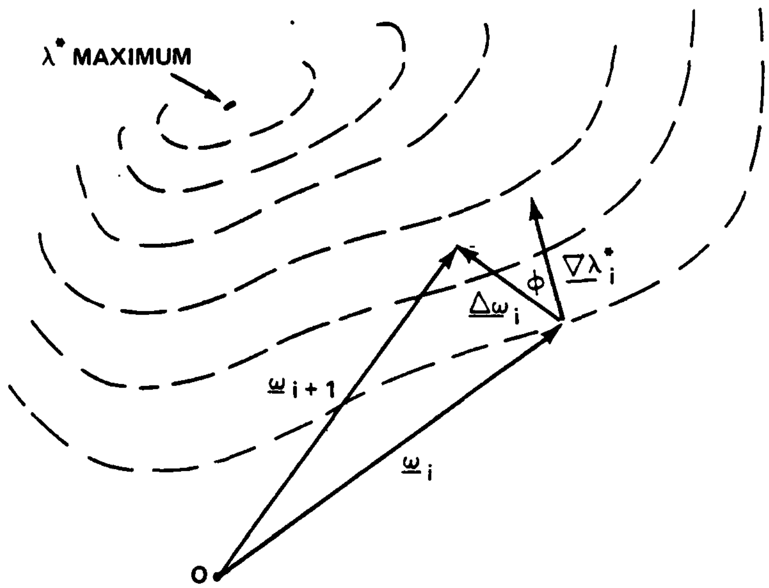


Figure 6. λ* contours and parameter vectors associated with the Newton-Raphson and method of scoring optimization schemes

---

[13]See footnote 13 on page 3-2.

To prevent the step $\underline{\Delta\omega}_i$ from overshooting the local region of increasing λ^* the step is often written in the modified form

$$\underline{\Delta\omega_i} = -\rho_i[E(\mathcal{JC})]^{-1}\ \underline{\nabla\lambda_i^*}\quad , \tag{4-8}$$

where $0 \leq \rho_i \leq 1$. The value of ρ_i can be chosen in various ways (e.g. see Fox, Reference 25) but Jennrich and Sampson (Reference 13) simply halve the preceeding step until an increased value of λ^* is obtained.

4-2 CONSTRAINED OPTIMIZATION

It is expected that for most sets of data the unconstrained N-R and Fisher scoring algorithms will produce estimates of $\underline{\omega}$ that lie in the feasible region of $\underline{\omega}$ space -- that is, the region corresponding to positive definite variance covariance matrices. When such is not the case, however, we need a method by which the maximum of λ^* within or at the boundary of the feasible region can be found. A popular and general method for effecting such a solution is referred to as the interior penalty function technique (see e.g. Reference 25 and 26). We will use a version of this method proposed by Carroll (Reference 27) that maximizes the function $\Lambda^*(\underline{\omega})$ instead of λ^*, where Λ^* is defined as

---

[25]Fox, R. L., Optimization Methods for Engineering Design (Reading, MA: Addison-Wesley Publishing Co., 1973).

[13]See footnote 13 on page 3-2.

[26]Aoki, M., Introduction to Optimization Techniques (New York: The Macmillan Co., 1971).

[27]Carroll, C. W., "The Created Response Surface Technique for Optimizing Nonlinear, Restrained Systems," Operations Research, Vol. 9, 169.

$$\Lambda^*(\underline{\omega}) = \lambda^*(\underline{\omega}) - \zeta(\underline{\omega}) \tag{4-9}$$

and
$$\zeta(\underline{\omega}) \equiv \sum_{\ell} \kappa_\ell / r_\ell(\underline{\omega}) \quad . \tag{4-10}$$

Here $\underline{\kappa} = \{\kappa_\ell\}$ is a set of positive constants and we assume that the set of constraint relationships can be expressed as $\{r_\ell(\underline{\omega})>0\}$. For a given starting point $\underline{\omega}_o$ within the feasible region the iteration process is conducted in the same manner as in the unconstrained problem. However, the iterate is now deflected away from the constraint boundaries by the penalty term $\zeta(\underline{\omega})$. When the process has converged to the maximum of Λ^* (within the feasible region) the process is restarted from the point of convergence using a smaller set of κ values in the objective function Λ^*. This sequence of operations is repeated until there is no appreciable change in the final Λ^* values. It should be noted that only those constraints in danger of being violated need be included in (4-10). That is, we might set a number of the κ coefficients to zero throughout the course of the computations.

In the present problem we require that the variance covariance matrices Γ and T be positive definite and $\sigma^2>0$. As Γ, given by (2-17), is diagonal it will be positive definite if and only if its elements are positive. And T, given by (2-19), will be positive definite if and only if the diagonal blocks Θ_i, $i = 1,2,\ldots,C$ are positive definite. It is convenient to specify the constraints on the $\{\Theta_i\}$ in terms of their discriminants. The mth discriminant of Θ_i, denoted as δ_{im}, is defined as the determinant of the upper left hand submatrix of size m. (This could also be called the upper left

hand principal minor of order m.) It can be shown (see Hildebrand, Reference 28, p. 51) that the p x p matrix Θ_i is positive definite if and only if $\delta_{im} > 0$, m = 1, 2,...,p. Consequently, we may now specify the required constraints in a manner that is consistent with the r notation used in (4-10). These are

$$\sigma^2 > 0, \tag{4-11}$$

$$\gamma_k > 0, \quad k = 1,2,\ldots,K \tag{4-12}$$

$$\delta_{im} > 0; \quad i = 1,2,\ldots,C, \; m = 1,2,\ldots,p \quad . \tag{4-13}$$

In order to use the unconstrained techniques developed earlier in this section for the purpose of maximizing Λ^* we need first and second partial derivatives of $\zeta(\underline{\omega})$. These are then subtracted from the corresponding derivatives of λ^* (or $\lambda^{\#}$) to form the derivatives of Λ^* (or $\Lambda^{\#}$). Making a self-explanatory change in the κ notation and defining c_{itm} as the cofactor of the element τ_{it} in the matrix with determinant δ_{im}, we can write the derivatives as

$$\frac{\partial \zeta}{\partial \sigma^2} = -\kappa(\sigma^2)/\sigma^4 \tag{4-14}$$

$$\frac{\partial \zeta}{\partial \gamma_k} = -\kappa(\gamma_k)/\gamma_k^2 \tag{4-15}$$

[28] Hildebrand, F. B., Methods of Applied Mathematics (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1952).

$$\frac{\partial \zeta}{\partial \tau_{it}} = - \sum_{m=1}^{p} \kappa(\delta_{im})\, c_{itm} / \delta_{im}^2 \tag{4-16}$$ *

$$\frac{\partial^2 \zeta}{\partial \sigma^2 \partial \sigma^2} = 2\, \kappa(\sigma^2)/\sigma^6 \tag{4-17}$$

$$\frac{\partial^2 \zeta}{\partial \gamma_k \partial \gamma_k} = 2\, \kappa(\gamma_k)/\gamma_k^3 \tag{4-18}$$

$$\frac{\partial^2 \zeta}{\partial \tau_{it} \partial \tau_{it}} = 2 \sum_{m=1}^{p} \kappa(\delta_{im})\, c_{itm}^2 / \delta_{im}^3 \tag{4-19}$$

$$\frac{\partial^2 \zeta}{\partial \tau_{is} \partial \tau_{it}} = \sum_{m=1}^{p} \kappa(\delta_{im}) \left[2 c_{itm}\, c_{ism} / \delta_{im}^3 - \frac{\partial c_{itm}}{\partial \tau_{is}} / \delta_{im}^2 \right] \tag{4-20}$$

$$\frac{\partial^2 \zeta}{\partial \gamma_k \partial \sigma^2} = \frac{\partial^2 \zeta}{\partial \tau_{it} \partial \sigma^2} = \frac{\partial^2 \zeta}{\partial \gamma_k \partial \gamma_\ell} = \frac{\partial^2 \zeta}{\partial \gamma_k \partial \tau_{it}} = \frac{\partial^2 \zeta}{\partial \tau_{hs} \partial \tau_{it}} = 0 \tag{4-21}$$

In (4-21) $k \neq \ell$ and $h \neq i$. For a value of p equal to 2 or 3 the value of $\partial c_{itm}/\partial \tau_{is}$ in (4-20) is 0 or ± 1 respectively with the signs of the latter depending on the detailed descriptions of the discriminant functions. We define a cofactor c_{itm} as zero if the δ_{im} discriminant function does not contain τ_{it}.

*Derivatives of a determinant $|A|$ may be obtained from its expansion in terms of cofactors $|A| = \Sigma_i\; a_{ij}\, c_{ij}$, where c_{ij} is the cofactor of a_{ij} in A. Hence, $\partial |A| / \partial a_{ij} = c_{ij}$ (see Searle, Reference 20, p. 86).

[20]See footnote 20 on page 3-10.

4-3 OTHER COMPUTATIONAL CONSIDERATIONS

Monitoring the progress of the iteration process in either the unconstrained or constrained settings will require the evaluation of λ^* or $\lambda^\#$. For this purpose we write them as

$$\lambda^* = -\frac{N}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log(|H|) - \frac{1}{2\sigma^2}\,\underline{y}'(\underline{y}-X\underline{\mu}-Z\underline{\nu}) \tag{4-22}$$

$$\lambda^\# = -\frac{(N-Cp)}{2}\log\sigma^2 - \frac{1}{2}\log(|H|\cdot|X'H^{-1}X|) - \frac{1}{2\sigma^2}\,\underline{y}'(y-X\underline{\mu}-Z\underline{\nu}), \tag{4-23}$$

which, with the exception of the determinants, involve quantities that are easily computed from previous results.

We can obtain the determinants $|H|$ and $|X'H^{-1}X|$ in forms more suitable for computations by using the identities (see Searle, Reference 20, p. 96)

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \equiv |G_{11}| \cdot |G_{22} - G_{21}G_{11}^{-1}G_{12}| \tag{4-24}$$

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \equiv |G_{22}| \cdot |G_{11} - G_{12}G_{22}^{-1}G_{21}| \quad , \tag{4-25}$$

which respectively assume the nonsingularity of G_{11} and G_{22}. By applying these identities to expand the determinant

[20] See footnote 20 on page 3-10.

$$\begin{vmatrix} (D^{-1} + Z'Z) & Z' \\ Z & I \end{vmatrix}$$

we find, as shown by Hemmerle and Hartley (Reference 11), that

$$|H| \equiv |D^{-1}+Z'Z| \cdot |D| \equiv |B| \cdot |D| \; . \tag{4-26}$$

Now, using the partitioned form of B given in (3-20) and the above identities we can write $|B|$ in two forms

$$|B| = \begin{vmatrix} U'U+\Gamma^{-1} & U'W \\ W'U & W'W+T^{-1} \end{vmatrix} \equiv |U'U+\Gamma^{-1}| \cdot |W'W+T^{-1}-W'U(U'U+\Gamma^{-1})^{-1}U'W| \tag{4-27}$$

and

$$|B| = \begin{vmatrix} U'U+\Gamma^{-1} & U'W \\ W'U & W'W+T^{-1} \end{vmatrix} \equiv |W'W+T^{-1}| \cdot |U'U+\Gamma^{-1} - U'W\,(W'W+T^{-1})^{-1}W'U| \; . \tag{4-28}$$

It is observed that the right most determinants in these equations involve matrices that are inverted to give $\overline{B}_{22}$ and $\overline{B}_{11}$ in (3-22a) and (3-24b). Since the determinant of a matrix can in many cases be obtained as a by-product of the inversion algorithm, e.g., a Cholesky decomposition (Reference 11), we can easily compute $|B|$ by either (4-27) or (4-28), depending upon which set of equations (3-22a) through (3-24a) or (3-22b) through (3-24b) are used to invert B. The other determinants needed to compute H are also easily obtained; $|U'U+\Gamma^{-1}|$ is the product of diagonal elements and $|D|$ and $|W'W+T^{-1}|$ are given by the products of the determinants of the diagonal blocks. To obtain $|X'H^{-1}X|$ we note that $\overline{A}_{11}$ given by (3-15) may be expressed as

[11]See footnote 11 on page 3-1.

$$\bar{A}_{11} = \{X'[I-ZD(I+Z'ZD)^{-1}Z']X\}^{-1} = [X'H^{-1}X]^{-1}, \tag{4-29}$$

from (3-7), (3-18), and (3-39). Hence, $|X'H^{-1}X|$ may be formed as a by-product of the inversion process by which we obtain $\bar{A}_{11}$.

As a final note concerning the numerical methods it is observed that many of the matrix operations required to invert matrices A and B and to calculate the derivatives can be done once and reused in the subsequent iterations. For example the products $X'X$, $X'U$, $X'W$, $U'U$, $U'W$, $W'W$, $\underline{y}'X$, $\underline{y}'U$, $\underline{y}'W$, and $\underline{y}'\underline{y}$ all involve known fixed constants. Furthermore, since the various equations used in the calculations depend on the larger (N rows) matrices X, U, W, and $\underline{y}$ through these products only, storage can be economized by computing the product matrices in advance and not saving X, U, W, and $\underline{y}$. It is also pointed out that in computing the trace of a matrix product, an operation that frequently appears in the expressions for the derivatives, it is only necessary to compute the diagonal elements of the product. For this reason and because the matrices appearing in these traces are generally of modest sizes, the computation of the derivatives should be rapid in comparison with the inversion times for matrices A and B.

CHAPTER 5

APPLICATIONS OF THE MODEL

In this section we indicate some of the applications for which the model may be used. No attempt is made to be exhaustive in this effort; rather, we will attempt to show by a few useful examples that the applications are wide ranging.

5-1 DIRECT RESULTS

Results obtained directly from the model computations are the ML or REML estimates of the following quantities (again we show just the ML estimates): the general experimental error variance σ^2; the gage calibration error variances for different gage classes given by

$$\hat{\sigma}^2_k \equiv \widehat{\mathrm{Var}(\alpha^*_{km})} = \hat{\gamma}_k \hat{\sigma}^2 \quad , \quad k = 1,2,\ldots, K; \tag{5-1}$$

the variances and covariances of the performance variations of different explosive classes given by

$$\hat{\Sigma}_i \equiv \widehat{\mathrm{Var}\ (\beta^*_{ij})} = \hat{\Theta}_i \hat{\sigma}^2 \quad , \ i = 1,2,\ldots, C \ ; \tag{5-2}$$

the similitude equation coefficients for different explosive classes given by

$$\underline{\hat{\mu}} = [\bar{A}_{11} X' + \bar{A}_{12} Z'] \underline{y} \ ; \tag{5-3}$$

the sample values of the gage class calibration errors $\underline{a}$ and explosive performance variations $\underline{b}$ given by

$$\begin{bmatrix} \hat{\underline{a}} \\ \hat{\underline{b}} \end{bmatrix} = [\bar{A}'_{12}X' + \bar{A}_{22}Z']\underline{y} \; ; \qquad (5\text{-}4)$$

and the error vector or vector of residuals given by

$$\hat{\underline{e}} = \underline{y} - X\hat{\underline{u}} - U\hat{\underline{a}} - W\hat{\underline{b}} \; . \qquad (5\text{-}5)$$

In these expressions and below the submatrices of A^{-1} are computed using the ML or REML estimates of the variance covariance matrices.

It is common in ordinary regression models to examine the model residuals for systematic trends that would indicate inconsistencies between the data and the model. Various techniques in this regard are discussed by Draper and Smith (Reference 29) and by Seber (Reference 7). In the present mixed model such examinations can be extended to include the estimates of all the random effects. Tests of the distributional assumptions can be made directly on the values of $\hat{\underline{a}}$, $\hat{\underline{b}}$, and $\hat{\underline{e}}$ using, for example, the methods summarized in Mehrotra and Michalek (Reference 30). The identification of faulty gages or of anomalously performing explosive charges should be immediately evident upon examination of $\hat{\underline{a}}$ and $\hat{\underline{b}}$.

---

[29]Draper, N. R., and Smith, H., Applied Regression Analysis (New York: John Wiley & Sons, 1966).

[7]See footnote 7 on page 2-6.

[30]Mehrotra, K. G., and Michalek, J. E., "Tests for Univariate and Multivariate Normality," RADC-TR-76-140, May 1976.

5-2 VARIANCE COVARIANCE MATRIX FOR $\hat{\underline{\mu}}$, $\hat{\underline{a}}$, AND $\hat{\underline{b}}$

From (3-12) $\underline{\mu}$, and $\underline{v} \equiv (\underline{a}', \underline{b}')'$ may be written as

$$\begin{bmatrix} \underline{\mu} \\ \underline{v} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \underline{y},$$

where

$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = A^{-1} \begin{bmatrix} X' \\ Z' \end{bmatrix} = \begin{bmatrix} \bar{A}_{11}X' + \bar{A}_{12}Z' \\ \bar{A}'_{12}X' + \bar{A}_{22}Z' \end{bmatrix} . \tag{5-7}$$

Here the components of A^{-1} are evaluated from the ML or REML estimates of the variance covariance component ratios $\underline{\theta}$ and σ^2. Since A^{-1} and $[R_1', R_2']'$ are complicated functions of $\underline{y}$ there is no known exact expression for the variance covariance matrix of $\hat{\underline{\mu}}$, $\hat{\underline{a}}$, and $\hat{\underline{b}}$. Nevertheless, it is a common practice in this situation (e.g., see References 14, p. 205 and 15, p. 35) to obtain an approximate variance covariance matrix by treating A^{-1} as though it were calculated from the true fixed values of $\underline{\theta}$. Equivalently, this presumes knowledge of the true value of H. Equation (5-6) then becomes a linear function of $\underline{y}$ and the variance covariance matrix is easily written down as (see footnote on page 2-6).

$$\mathrm{Var} \begin{bmatrix} \hat{\underline{\mu}} \\ \hat{\underline{v}} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} (I+ZDZ') \begin{bmatrix} R_1', R_2' \end{bmatrix} \sigma^2$$

$$= \begin{bmatrix} R_1(I+ZDZ')R_1' & R_1(I+ZDZ')R_2' \\ R_2(I+ZDZ')R_1' & R_2(I+ZDZ')R_2' \end{bmatrix} \sigma^2. \tag{5-8}$$

---

[14] See footnote 14 on page 3-2.

[15] See footnote 15 on page 3-2.

Following Henderson (Reference 31, Appendix A) this can be simplified by use of the relation

$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} [X,Z] = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}'_{12} & \bar{A}_{22} \end{bmatrix} \left\{ \begin{bmatrix} X'X & X'Z \\ Z'X & D^{-1}+Z'Z \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & D^{-1} \end{bmatrix} \right\}$$

$$= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} 0 & \bar{A}_{12}D^{-1} \\ 0 & \bar{A}_{22}D^{-1} \end{bmatrix} . \tag{5-9}$$

Hence,

$$\begin{aligned} R_1X &= I \\ R_1Z &= -\bar{A}_{12} \\ R_2X &= 0 \\ R_2Z &= I - \bar{A}_{22}D^{-1} . \end{aligned} \tag{5-10}$$

By substituting (5-7) and (5-10) into (5-8) one readily obtains

$$\mathrm{Var} \begin{bmatrix} \underline{\mu} \\ \underline{v} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & D-A_{22} \end{bmatrix} \sigma^2 . \tag{5-11}$$

An estimate of (5-11) can be obtained by substituting the ML or REML estimates of D, $\bar{A}_{11}$, $\bar{A}_{12}$, and σ^2. The result $\mathrm{Var}(\underline{\mu}) = \bar{A}_{11}\sigma^2$ could have also been derived as done by Corbeil and Searle (Reference 15) from (3-5) and (4-29).

In Reference 14 (p. 205) Harville has pointed out that (5-11), because H is assumed to be known, will tend to underpredict the dispersions of $\hat{\underline{\mu}}$ and

[31] Henderson, C. R., "Best Liner Unbiased Estimation and Prediction Under a Selection Model," Biometrics, Vol. 31, p. 423.

[15] See footnote 15 on page 3-2.

[14] See footnote 14 on page 3-2.

$(\hat{\underline{a}}', \hat{\underline{b}}')'$ but suggests also that "the downward bias may, at least in some instances, be so small as to be unimportant." Quantification of the accuracy of this result (as well as of other results derived below in a similar manner) is possible by means of Monte Carlo simulation studies. Such studies will be pursued upon completion of the model coding and published in a subsequent report.

5-3 LARGE SAMPLE VARIANCE COVARIANCE MATRICES FOR ESTIMATORS OF COMPONENTS OF VARIANCE AND COMPONENTS OF VARIANCE RATIOS

A well known result from the theory of maximum likelihood (e.g., see Kendall and Stuart, Reference 32) is that the asymptotic ($N \to \infty$) variance covariance matrix of the vector of parameter estimates is given by the inverse of the information matrix $\mathcal{J}$, which is the negative of the matrix of expected values of the second derivatives of the log-likelihood. In our case with (4-1) as the vector of parameters we have

$$\mathcal{J} = -E(\mathcal{H}), \qquad (5\text{-}12)$$

where $\mathcal{H}$ is given by (4-4). Hence, denoting the vector of ML estimates as

$$\hat{\underline{\omega}} \equiv (\hat{\sigma}^2, \hat{\gamma}_1, \ldots, \hat{\gamma}_K, \hat{\tau}_{11}, \ldots \hat{\tau}_{1\pi}, \ldots, \hat{\tau}_{C1}, \ldots, \hat{\tau}_{C\pi})' \qquad (5\text{-}13)$$

we can write

$$\mathrm{Var}(\hat{\underline{\omega}}) \geq \mathcal{J}^{-1}, \qquad (5\text{-}14)$$

[32]Kendall, M. G., and Stuart, A., The Advanced Theory of Statistics (London: Charles Griffin & Co., Ltd., 1973), Vol. 2.

where equality is achieved in the limit as $N\to\infty$. This lower bound is, however, often used as an approximation for $\text{Var}(\hat{\underline{\omega}})$. An estimate of $\text{Var}(\hat{\underline{\omega}})$ may thus be obtained from the final iterated value of $E(\mathcal{H})$.

Often it may be of more interest to know the variance covariance matrix of the estimators of the components of variance rather than of the components of variance ratios. Suppose this is denoted as

$$\hat{\underline{\varphi}} \equiv (\hat{\sigma}^2, \hat{\sigma}^2_1, \ldots, \hat{\sigma}^2_K, \hat{\upsilon}_{11}, \ldots \hat{\upsilon}_{1\pi}, \ldots, \hat{\upsilon}_{C1}, \ldots, \hat{\upsilon}_{C\pi})' \quad , \tag{5-15}$$

where as in (5-1) $\hat{\sigma}^2_k \equiv \hat{\gamma}_k \hat{\sigma}^2$, $k = 1, \ldots, K$ and $\hat{\upsilon}_{ij} \equiv \hat{\tau}_{ij} \hat{\sigma}^2, i = 1, \ldots, C,\ j = 1, \ldots, \pi$. Corbeil and Searle (Reference 15) have shown that the relationship between $\text{Var}(\hat{\underline{\varphi}})$ and $\text{Var}(\hat{\underline{\omega}})$ is given by

$$\text{Var}(\hat{\underline{\varphi}}) = \Omega \, \text{Var}(\hat{\underline{\omega}}) \, \Omega' \quad , \tag{5-16}$$

where Ω is the Jacobian matrix for the transformation from $\underline{\varphi}$ to $\underline{\omega}$. This is easily shown to be

$$\Omega = \begin{bmatrix} 1 & 0 \\ \underline{\omega}_- & \sigma^2 I \end{bmatrix} , \tag{5-17}$$

where $\underline{\omega}_-$ is defined by $\underline{\omega} \equiv (\sigma^2, \underline{\omega}_-')'$. Hence, the large sample value of $\text{Var}(\hat{\underline{\varphi}})$ can be obtained by substituting J^{-1} for $\text{Var}(\hat{\underline{\omega}})$ in (5-16).

[15]See footnote 15 on page 3-2.

5-4 PREDICTION OF THE PERFORMANCE IN A PAST TEST

In our discussion of the basic model for a single observation, Section 2-1, we let $f_{ij}(x)$ denote the value of the principal quantity of interest in the shot of the ith explosive class at some reduced travel distance x from the charge. We assume $f_{ij}(x)$ is any 1 to 1 transformation, such as the logarithm, of a possibly reduced performance index or effectiveness factor that has physical meaning and significance. In the light of past discussions we will now use a somewhat more explicit notation and let $f_{ij}(x)$ denote the realization of the random variable $f^*_{ij}(x)$ that in accord with (2-7) may be expressed as

$$f^*_{ij}(x) = \underline{\Phi}'\underline{\mu}_i + \underline{\Phi}'\underline{\beta}^*_{ij} \ . \tag{5-18}$$

Here $\underline{\Phi}$ denotes some vector function of the arbitrary distance x analogous to $\underline{\Phi}_{ijkm}$ described earlier. Following suit, we write the unknown realization of $f^*_{ij}(x)$, that is the sought after quantity, as

$$f_{ij}(x) = \underline{\Phi}'\underline{\mu}_i + \underline{\Phi}'\underline{\beta}_{ij} \ , \tag{5-19}$$

and the ML estimator of $f_{ij}(x)$ as

$$\hat{f}_{ij}(x) = \underline{\Phi}'\hat{\underline{\mu}}_i + \underline{\Phi}'\hat{\underline{\beta}}_{ij} \ . \tag{5-20}$$

Here $\hat{\underline{\beta}}_{ij}$ and $\underline{\beta}_{ij}$ are the subvectors of the vectors $\hat{\underline{b}}$ and $\underline{b}$ that correspond to the jth shot of the ith explosive class.

Confidence limits for $f_{ij}(x)$ may be obtained under the assumption of a known variance covariance ratios matrix H. The limits so obtained may be reasonably accurate when sample sizes are large, but when samples are only modestly sized the confidence interval will correspond to a confidence coefficient that is smaller than the one specified. Nevertheless, it may be possible to attach a more realistic confidence coefficient to the interval by a Monte Carlo simulation procedure. Hence, we include the theory in this report.

Dropping the argument indicating the explicit dependency on x, we can define a vector $\underline{\Psi}$ such that

$$f_{ij} = \underline{\Psi}' \begin{bmatrix} \underline{\mu} \\ \underline{v} \end{bmatrix} . \tag{5-21}$$

Comparing this with (5-19) it is obvious that $\underline{\Psi}$ consists of a column of zeros imbeded with two $\underline{\Phi}$ vectors located in such a way as to extract the $\underline{\mu}_i$ and $\underline{\beta}_{ij}$ subvectors from $\underline{\mu}$ and $\underline{v}$. Similarly we can write

$$\overset{*}{f}_{ij} = \underline{\Psi}' \begin{bmatrix} \underline{\mu} \\ \underline{v}^* \end{bmatrix} \tag{5-22}$$

and

$$\hat{f}_{ij} = \underline{\Psi}' \begin{bmatrix} \hat{\underline{\mu}} \\ \hat{\underline{v}} \end{bmatrix} . \tag{5-23}$$

The development of confidence limits for f_{ij} can proceed in a manner similar to that found in a more general study due to Harville (Reference 33). Under the assumption of known variance covariance ratios, $\hat{\underline{\mu}}$ is unbiased and it

[33] Harville, D. A., "Confidence Intervals and Sets for Linear Combination of Fixed and Random Effects," Biometrics, Vol. 32, 1976, p. 403.

can be shown (see Henderson, Reference 31 for an explicit derivation) that the random variable $\hat{f}_{ij}-f^*_{ij}$ is normally distributed with a zero mean and variance $\underline{\Psi}'A^{-1}\underline{\Psi}\sigma^2$; hence

$$(\hat{f}_{ij}-f^*_{ij})/(\underline{\Psi}'A^{-1}\underline{\Psi}\sigma^2)^{1/2} \sim N(0,1) \quad . \tag{5-24}$$

Furthermore,

$$\tilde{\sigma}^2 = \underline{y}'\underline{\hat{e}}/(N-Cp) \tag{5-25}$$

is an unbiased estimator of σ^2 and it can be shown that $(N-CP)\tilde{\sigma}^2/\sigma^2$ is χ^2 distributed with (N-Cp) degrees of freedom and is independent of (5-23). From these results it follows that

$$\frac{\hat{f}_{ij}- f^*_{ij}}{(\underline{\Psi}'A^{-1}\underline{\Psi})^{1/2}\tilde{\sigma}} \sim t_{N-Cp}, \tag{5-26}$$

i.e., has a t distribution with N-Cp degrees of freedom. Then, from the symmetry of the t distribution, we have

$$\Pr[\,|\hat{f}_{ij}-f^*_{ij}| \le (\underline{\Psi}'A^{-1}\underline{\Psi})^{1/2}\,\tilde{\sigma}\, t_{\delta/2,N-Cp}] = 1 - \delta\ , \tag{5-27}$$

where $t_{\delta/2,N-Cp}$ is the $100\delta/2$ percentile of the t_{N-Cp} distribution. Hence, for a particular realization we find that

$$f_{ij} \pm (\underline{\Psi}'A^{-1}\underline{\Psi})^{1/2}\,\tilde{\sigma}\, t_{\delta/2,N-Cp} \tag{5-28}$$

[31]See footnote 31 on page 5-4.

are 1-δ confidence limits for f_{ij} (termed "unconditional" confidence limits by Harville in Reference 33). These, of course, are valid for any values of the regressor variable x.

For computation purposes it is useful to express $\underline{\psi}' A^{-1} \underline{\psi}$ in the following manner

$$\underline{\psi}' A^{-1} \underline{\psi} = \underline{\phi}' \overline{A}_{11}(\underline{\mu}_i) \underline{\phi} + 2 \underline{\phi}' \overline{A}_{12}(\underline{\mu}_i, \underline{\beta}_{ij}) \underline{\phi} + \underline{\phi}' \overline{A}_{22}(\underline{\beta}_{ij}) \underline{\phi} . \tag{5-29}$$

Here $\overline{A}_{11}(\underline{\mu}_i)$ and $\overline{A}_{22}(\underline{\beta}_{ij})$ denote the p x p submatrices of A^{-1} having rows and columns corresponding to $\underline{\mu}_i$ and $\underline{\beta}_{ij}$ respectively and the submatrix $\overline{A}_{12}(\underline{\mu}_i, \underline{\beta}_{ij})$ corresponds to the $\underline{\mu}_i$ rows and the $\underline{\beta}_{ij}$ columns.

If a logarithmic transformation of the response variable has been employed it may be of interest to estimate and obtain confidence limits for the antilog of f_{ij}. Suppose the common logarithm was used. Then the ML estimator of the (possibly reduced) performance index T_{ij} in the jth shot of the ith explosive class would be

$$\hat{\underline{T}}_{ij} = 10^{\hat{f}_{ij}} \tag{5-30}$$

and from (5-27) confidence limits for T_{ij} would be given by

$$10^{[\hat{f}_{ij} \pm (\underline{\psi}' A^{-1} \underline{\psi})^{1/2} \tilde{\sigma}\, t_{\delta/2, N-Cp}]} . \tag{5-31}$$

[33]See footnote 33 on page 5-8.

5-5 PREDICTION OF THE PERFORMANCE IN A FUTURE TEST

Continuing the notation of the previous section, we now consider the prediction of the performance index in a future test at an arbitrary distance x from the charge. We assume the charge is of the ith explosive class where $1 \leq i \leq C$. To avoid confusion with past quantities we will indicate the future value by f_{i*}, i.e., an asterick in place of the subscript denoting a particular shot.

In analogy with (5-18) we let f_{i*} be the realization of the random variable

$$f^*_{i*} = \underline{\phi}'\underline{\mu}_i + \underline{\phi}'\underline{\beta}^*_{i*} \quad , \tag{5-32}$$

where β^*_{i*} is independent of past quantities and therefore of $\underline{y}$ and $\hat{\underline{u}}$. Under the assumption of a known H, an unbiased estimator of f_{i*} is given by

$$\hat{f}_{i*} = \underline{\phi}'\hat{\underline{\mu}}_i \; . \tag{5-33}$$

To obtain confidence limits for f_{i*} under the same assumption consider the random variable $\hat{f}_{i*} - f^*_{i*}$. We have

$$E\,(\hat{f}_{i*} - f^*_{i*}) = \underline{\phi}'\underline{\mu}_i - \underline{\phi}'\underline{\mu}_i = 0 \tag{5-34}$$

and

$$\mathrm{Var}(\hat{f}_{i*} - f^*_{i*}) = \underline{\phi}'(\bar{A}_{11}(\underline{\mu}_i) + \Theta_i)\,\underline{\phi}\sigma^2 \; , \tag{5-35}$$

so that

$$\hat{f}_{i*} - f^*_{i*} \sim N(0,\; \underline{\phi}'(A_{11}(\underline{\mu}_i) + \Theta_i)\,\underline{\phi}\sigma^2). \tag{5-36}$$

Now, since $\tilde{\sigma}^2$ of (5-24) is also independent of f^*_{i*}, we have

$$\frac{\hat{f}_{i*}-f^*_{i*}}{[\underline{\phi}'(A_{11}(\mu_i) + \Theta_i)\underline{\phi}]^{1/2}\tilde{\sigma}} \sim t_{N-Cp} \tag{5-37}$$

This allows us to make the probability statement

$$\Pr\left\{|\hat{f}_{i*} - f^*_{i*}| \leq [\underline{\phi}'(A_{11}(\underline{\mu_i}) + \Theta_i)\underline{\phi}]^{1/2}\, \tilde{\sigma}\, t_{\delta/2,N-Cp}\right\} = 1 = \delta\ . \tag{5-38}$$

Hence, 1-δ confidence limits for a particular future realization f_{i*} are given by

$$f_{i*} \pm [\underline{\phi}'(A_{11}(\underline{\mu_i}) + \Theta_i)\underline{\phi}]^{1/2}\, \tilde{\sigma}\ \ t_{\delta/2,N-Cp}\ \ . \tag{5-39}$$

As emphasized in the previous section, these limits will tend to underpredict the size of the confidence interval because we have assumed the estimated value of H to be the true value.

If a common logarithmic transformation of the data has been used, the ML estimate of $T_{i*} = 10^{f_{i*}}$ can be obtained from

$$\hat{T}_{i*} = 10^{\hat{f}_{i*}}\quad , \tag{5-40}$$

and approximate confidence limits for T_{i*} from

$$10^{\{f_{i*} \pm [\phi'(\bar{A}_{11}(\underline{\mu_i}) + \Theta_i)\phi]^{1/2}\, \tilde{\sigma}\, t_{\delta/2,N-Cp}\}}\ . \tag{5-41}$$

BIBLIOGRAPHY

Aoki, M., Introduction to Optimization Techniques (New York: The Macmillan Co., 1971).

Bartlett, M. S., "An Inverse Matrix Adjustment Arising in Discriminant Analysis," Annals of Mathematical Statistics, Vol. 22, 1951, p. 107.

Brown, R. H., "Analysis of Data When Several Sources of Variation are Present," Explosives Research Memorandum 22, Navy Dept. Bureau of Ordnance, 1 Dec 1944.

Carroll, C. W., "The Created Response Surface Technique for Optimizing Nonlinear, Restrained Systems," Operations Research, Vol. 9, 169.

Cole, R. H., Underwater Explosions (New Jersey: Princeton University Press, 1948).

Corbeil, R. R., and Searle, S. R., "A Comparison of Variance Component Estimators," Biometrics, Vol. 32, 1976, p. 779.

Corbeil, R. R., and Searle, S. R., "Restricted Maximum Likelihood (REML) Estimation of Variance Components in the Mixed Model," Technometrics, Vol. 18, No 1, 1976, p. 31.

BIBLIOGRAPHY (Cont.)

Draper, N. R., and Smith, H., Applied Regression Analysis (New York: John Wiley & Sons, 1966).

Fox, R. L., Optimization Methods for Engineering Design (Reading, MA: Addison-Wesley Publishing Co., 1973).

Gaspin, J. B., "Validation of a Gage Location Method for Underwater Explosion Tests," NSWC TR in preparation.

Goertner, J. F., "Scaling Underwater Explosion Shock Waves for Differences in Ambient Sound Speed and Density," NSWC TR 80-491, 18 Dec 1980.

Hartley, H. O., and Rao, J. N. K., "Maximum-Likelihood Estimation for the Mixed Analysis of Variance Model," Biometrika, Vol. 54, 1 and 2, 1967, p. 93.

Harville, D. A., "Bayesian Inference for Variance Components Using Only Error Contrasts," Biometrika, Vol. 61, 1974, p. 383.

Harville, D. A., "Confidence Intervals and Sets for Linear Combinations of Fixed and Random Effects," Biometrics, Vol. 32, 1976, p. 403.

Harville, D. A., "Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects," The Annals of Statistics, Vol. 4, No. 2, 1976, p. 384.

BIBLIOGRAPHY (Cont.)

Harville, D. A., "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," Journal of the American Statistical Assoc., Vol. 72, No. 358, 1977, p. 320.

Harville, D. A., "Some Useful Representations for Constrained Mixed-Model Estimation," Journal of the American Statistical Assoc., Vol. 74, No. 365, 1979, p. 200.

Hemmerle, W. J., and Hartley, H. O., "Computing Maximum Likelihood Estimates for the Mixed A. O. V. Model Using the W Transformation," Technometrics, Vol. 15, No. 4, 1973, p. 819.

Hemmerle, W. J., and Lorens, J. A., "Improved Algorithm for the W-Transform in Variance Component Estimation," Technometrics, Vol. 18, No. 2, 1976, p. 207.

Henderson, C. R., "Best Linear Unbiased Estimation and Prediction Under a Selection Model," Biometrics, Vol. 31, p. 423.

Henderson, C. R., Kempthorne, O., Searle, S. R., and Von Krosigk, C. N., "Estimation of Environmental and Genetic Trends from Records Subject to Culling," Biometrics, Vol. 15, 1959, p. 192.

Hildebrand, F. B., Methods of Applied Mathematics (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1952).

BIBLIOGRAPHY (Cont.)

Jennrich, R. I., and Sampson, P. F., "Newton-Raphson and Related Algorithms for Maximum Likelihood Variance Component Estimation," Technometrics, Vol. 18, No. 1, 1976, p. 11.

Kendall, M. G., and Stuart, A., The Advanced Theory of Statistics (London: Charles Griffin & Co., Ltd., 1973) Vol. 2.

Mehrotra, K. G., and Michalek, J. E., "Tests for Univariate and Multivariate Normality," RADC-TR-76-140, May 1976.

Nering, E. D., Linear Algebra and Matrix Theory (New York: John Wiley & Sons, 2nd Ed, 1970).

Patterson, H. D., and Thompson, R., "Recovery of Inter-Block Information When Block Sizes are Unequal," Biometrika, Vol. 58, No. 3, 1971, p. 545.

Searle, S. R., Linear Models (New York: John Wiley & Sons, Inc., 1971).

Searle, S. R., Matrix Algebra for the Biological Sciences (New York: John Wiley & Sons, Inc., 1966).

Searle, S. R., "Topics in Variance Component Estimation," Biometrics, Vol. 27, 1971, p. 1.

BIBLIOGRAPHY (Cont.)

Seber, G. A. F., Linear Regression Analysis (New York: John Wiley & Sons, Inc., 1977).

Snay, H. G., "Model Tests and Scaling," NOLTR 63-257, 1 Dec 1964.

Westlake, J. R., A Handbook of Numerical Matrix Inversion and Solution of Linear Equations (Huntington, NY: Robert E. Krieger Publishing Co., 1975).

DISTRIBUTION

	Copies
Chief of Naval Material	
Office of Naval Technology	
Attn: NMAT-07P (Dr. J. Enig)	1
Washington, DC 20360	
Chief of Naval Operations	
Department of the Navy	
Attn: OP-981F	1
OP-981F1	1
OP-324	1
OP-324T	1
Washington, DC 20350	
Chief of Naval Research	
800 N. Quincy Street	
Attn: Code 411 S&P	
(Dr. E. Wegman)	2
Code 411 OR (S. Selig)	1
Arlington, VA 22217	
Commander	
Naval Air Systems Command	
Department of the Navy	
Attn: NAIR-350D	1
NAIR-310A (Dr. G. Heiche)	1
Washington, DC 20361	
Commander	
Naval Sea Systems Command	
Attn: SEA-63R (F. Romano)	1
SEA-X5511 (J. Sullivan)	1
SEA-X5511 (D. M. Hurt)	1
SEA-62R32 (G. D. Edwards)	1
SEA-62R (W. W. Blaine)	1
SEA-99612	1
SEA-32R (C. Pohler)	1
PMS-402	1
PMS-406	1
PMS-407	1
Washington, DC 20362	

DISTRIBUTION (Cont.)

	Copies
Director	
Naval Research Laboratory	
Attn: Technical Information Section	1
G. O'Hara	1
R. Bort	1
Washington, DC 20375	
Commander	
David Taylor Naval Research & Development Center	
Attn: Library	1
Code 17 (Dr. W. Murray)	1
Code 1740 (Dr. F. Fisch)	1
Code 175 (J. Sykes)	1
Code 175.3 (W. Conley)	1
Code 175.3 (P. Manny)	1
Bethesda, MD 20084	
David Taylor Naval Ship Research & Development Center	
Underwater Explosions Research Division	
Attn: Technical Library	1
Code 177 (R. Fuss)	1
Code 177.1 (V. Bloodgood)	1
Portsmouth, VA 23709	
Commander	
Naval Weapons Center	
Attn: Code 533 (Technical Library	1
Code 385 (A. B. Amster)	1
China Lake, CA 93555	
Superintendent	
Naval Academy	
Attn: Dr. J. Geremia	1
Annapolis, MD 21402	

DISTRIBUTION (Cont.)

	Copies
Superintendent Naval Postgraduate School Attn: Library Monterey, CA 93940	1
Commander Naval Coastal Systems Center Attn: Library Panama City, FL 32407	1
Commander Naval Ocean Systems Center Attn: Technical Library San Diego, CA 92152	1
Commanding Officer Naval Underwater Systems Center Attn: Technical Library Newport, RI 02840	1

	Copies
Commanding Officer Naval Weapons Station Attn: Dr. Lewis R. Rothstein (NEDED)	1
Dr. William McBride (NEDED)	1
Yorktown, VA 23491	
Director Waterways Experiment Station Attn: Technical Library	1
K. Davis	1
J. R. Britt	1
WESNE	1
P. O. Box 631 Vicksburg, MI 39180	
Commanding Officer Aberdeen Research and Development Center Attn: Technical Library	1
Dr. R. Moore	1
Aberdeen, MD 21005	

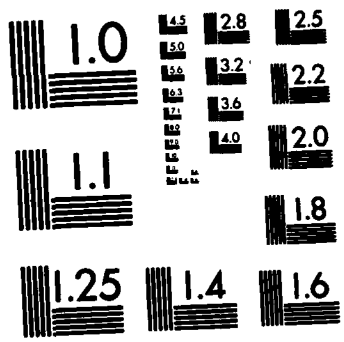
AD-A125 337   A MIXED REGRESSION MODEL FOR THE ANALYSIS OF EXPLOSIVES PERFORMANCE DATA(U) NAVAL SURFACE WEAPONS CENTER SILVER SPRING MD W W MCDONALD 30 JUL 82 NSWC/TR-82-74   2/2

UNCLASSIFIED   SBI-AD-F500 124   F/G 19/4   NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DISTRIBUTION (Cont.)

	Copies
Army Research Office Mathematics Division Attn: Dr. R. Launer Research Triangle Park, NC 27709	1
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12
Director Defense Advanced Research Project Agency Attn: Technical Library Washington, DC 20301	1
Defense Nuclear Agency Attn: SPSS Washington, DC 20305	2

	Copies
Iowa State University Dept. of Statistics Attn: Prof. D. Harville Ames, IA 50011	1
Cornell University Biometrics Unit Attn: Prof. S. Searle Ithaca, NY 14853	1
The George Washington University Dept. of Operations Research Attn: Prof. N. Singpurwalla Washington, DC 20052	1
Dr. Hans G. Snay 17613 Tree Lawn Drive Ashton, MD 20702	1

Internal Distribution	Copies
P. Wessel (R)	1
J. Proctor (R10)	1
I. Blatstein (R40)	1
L. Roslund (R122)	2
M. Marshall (R14)	1
R. Barash (R14)	1
T. Farley (R14)	1
J. Gaspin (R14)	1
J. A. Goertner (R14)	1
J. F. Goertner (R14)	1
A. Henney (R14)	1
N. Holland (R14)	1
D. Lehto (R14)	1
W. McDonald (R14)	20
M. Moussouros (R14)	1
D. Nicholson (R14)	1
E. Tanglis (R14)	1
R. Thrun (R14)	1
G. Young (R14)	1
M. Marcus (R14)	1
D. Phillips (R102)	1
E. Anderson (R12)	1
M. Stosz (R121)	1
F. Menz (R12)	1

	Copies
J. Pittman (R15)	1
M. Thomas (K106)	1
G. Gemmill (K106)	1
E. Cohen (R44)	1
R. Price (R15)	1
J. Ward (R15)	1
R. Tussing (R15)	1
R. Lorenz (R15)	1
J. Dempsey (R15)	1
W. Faux (R15)	1
J. Short (R12)	1
R. Bowen (R16)	1
F. Porzel (R15)	1
J. Dempsey (R15)	1
M. Swisdak (R15)	1
K. Mueller (R11)	1
C. Dickinson (R11)	1
T. Austin (R16)	1
H. Haiss (R17)	1
K. Bannister (R14)	1
E35	1
E341	9
E342	3